

UNIVERSITY OF CALIFORNIA, SAN DIEGO

**The turbulent and wavy upper ocean: transition from geostrophic flows to
internal waves and stimulated generation of near-inertial waves**

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy

in

Oceanography

by

Cesar Barbedo Rocha

Committee in charge:

Professor William R. Young, Chair
Professor Sarah T. Gille
Professor Teresa K. Chereskin
Professor Juan Carlos Del Álamo
Professor Stefan G. Llewellyn Smith

2018

Copyright
Cesar Barbedo Rocha, 2018
All rights reserved.

The dissertation of Cesar Barbedo Rocha is approved, and
it is acceptable in quality and form for publication on mi-
crofilm and electronically:

Chair

University of California, San Diego

2018

EPIGRAPH

Busca por el agrado de buscar, no por el de encontrar.

Jorge Luis Borges, *Fragmentos de un Evangelio Apócrifo*

TABLE OF CONTENTS

Signature Page	iii
Epigraph	iv
Table of Contents	v
List of Figures	vii
List of Tables	xiii
Acknowledgements	xiv
Abstract of the Dissertation	xvi
Chapter 1 Introduction	1
1.1 Background	1
1.2 Motivation	4
1.3 Dissertation's outline	5
Chapter 2 Mesoscale to submesoscale wavenumber spectra in Drake Passage	8
2.1 Introduction	8
2.2 Theoretical predictions	9
2.3 Velocity observations	10
2.4 Satellite altimetry	16
2.5 Some experiments with synthetic data	20
2.6 A high-resolution primitive-equation model	21
2.7 Discussion	27
2.8 Conclusions	34
Chapter 3 Seasonality of submesoscale dynamics in the Kuroshio Extension .	46
3.1 Introduction	46
3.2 The LLC numerical simulations	47
3.3 Bulk statistics of the surface lateral velocity gradient tensor .	49
3.4 Joint probability density functions	52
3.5 Wavenumber-frequency spectrum	53
3.6 Discussion and conclusion	55
Chapter 4 Stimulated generation: extraction of energy from balanced flow by near-inertial waves	58
4.1 Introduction	58
4.2 The vertical-plane-wave model	62
4.3 Conservation laws of the vertical-plane-wave model .	69
4.4 2D turbulence modified by near-inertial waves	76
4.5 Wave escape	85
4.6 The Eulerian-mean viewpoint	89
4.7 Discussion and conclusions	94

Chapter 5	Equilibration of forced 2D turbulence by stimulated generation	104
5.1	Introduction	104
5.2	The dissipative Boussinesq equations	107
5.3	The expansion	110
5.4	Local PV generation—an illustrative example	118
5.5	Energetics of the dissipative QG-NIW model	122
5.6	Forced-dissipative solutions	126
5.7	Conclusions	134
Chapter 6	On Galerkin approximations of the surface-active quasigeostrophic equations	139
6.1	Introduction	139
6.2	The exact system	141
6.3	Galerkin approximation using standard vertical modes	143
6.4	Three approximations	147
6.5	The Eady problem	153
6.6	The Green problem	158
6.7	Summary and conclusions	163

LIST OF FIGURES

Figure 1.1: Rotary frequency spectra of kinetic energy from WHOI299 mooring at 260 m (nominal depth). The mooring was located at (34.98°N, 152.03°E), near the Kuroshio Current; the local depth was 6149 m. The red line shows the spectrum of clockwise motions and the blue line shows the spectrum of anti-clockwise motions. The inertial period is $2\pi/f_0 \approx 20.88$ h.	2
Figure 1.2: (a) Near-inertial and subinertial velocity time series from WHOI299 mooring at 260 m (nominal depth), whose spectra is shown in figure 1.1. (b) A blow-up of the shaded timespan in (a). (c) Near-inertial velocity hodograph for the velocity series in (b); black dots in (c) show data on May 15, 1981, depicting the clockwise polarization of the near-inertial currents.	3
Figure 2.1: (a) Gridded mean (vectors) and eddy kinetic energy (colors) from ADCP data at 26 m (uppermost bin). The black line depicts the mean position of the Polar Front [1993-2013, updated from Salée, Speer, and Morrow 2008]. (b) Segments of tracks from AltiKa altimeter. (c) Subdomain of llc4320 simulation used in this study. Colors represent a snapshot of surface vertical vorticity normalized by the local inertial frequency.	11
Figure 2.2: ADCP KE spectra for three layers: (a) 26-50 m, (b) 58-98 m, and (c) 106-202 m; and wave-vortex decomposition for the 58-98 m layer: (d) total kinetic energy, (e) the geostrophic flow (vortex) and (f) the inertia-gravity wave components of across-track and along-track components. The figure depicts across-track (red line) and along-track (blue line) KE spectra, and the spectral decomposition into horizontally rotational (green line) and divergent components (yellow line). Shaded represent 95% confidence limits. For reference, k^{-2} and k^{-3} curves are plotted (gray lines).	17
Figure 2.3: ADCP KE spectra for the 58-98 m layer from sub-transects to the north (solid) and to the south (dashed) of the Polar Front.	18
Figure 2.4: AltiKa SSH variance along-track wavenumber spectra. For comparison, also shown is the equivalent spectrum inferred from the ADCP spectrum by assuming geostrophy.	20

Figure 2.5: Synthetic data KE spectra and its decompositions into rotational and divergent components for different combinations of horizontally rotational and divergent flows. (a) κ^{-3} rotational flow, (b) κ^{-3} rotational flow combined with a high-pass (> 40 km) κ^{-3} divergent flow, and (c) κ^{-3} rotational flow combined with a high-pass (> 40 km) κ^{-2} divergent flow. In (c), the energy level of the divergent flow is a quarter of the energy level of the rotational flow. For reference, k^{-2} and k^{-3} curves are plotted (gray lines).	22
Figure 2.6: Comparison between ADCP and model spectra for the 58-98 m layer. For reference, k^{-2} and k^{-3} curves are plotted (gray lines).	27
Figure 2.7: Model KE spectra at the surface. The figure depicts spectra for hourly (red), and the “ LMG sampled” and block-averaged fields (yellow and green line). For reference, k^{-2} and k^{-3} curves are plotted (gray lines).	28
Figure 2.8: Model along-track wavenumber KE spectra for the 25-60 m layer: (a) hourly, (b) daily-averaged; (c) comparison of KE spectra of daily-averaged fields for three layers (solid lines represent across-track spectra, dashed lines along-track spectra); and model isotropic KE spectra at the surface: (d) hourly and (e) daily-averaged; (f) comparison of residual spectrum with the Garrett-Munk spectrum at 100 m. Shades represent 95 % confidence limits. For reference, and k^{-2} and k^{-3} curves are plotted; in (f) the Garrett-Munk spectrum is plotted with 1.4 times the energy level in Munk [1981].	29
Figure 2.9: Isotropic wavenumber SSH variance spectrum. Shades represent 95% confidence limits. For reference, $k^{-3/2}$ and k^{-5} curves are plotted (gray lines).	30
Figure 2.10: Horizontal wavenumber frequency spectrum of KE at 100 m. The solid line is the dispersion relationship for inertia-gravity waves integrated over vertical wavenumber, and the dashed lines indicate tidal and inertial frequencies. The inertial frequency is based on the mean latitude of the domain (about 59°S).	30
Figure 2.11: Base state used in the linear stability calculations.	42
Figure 2.12: Growth rate in horizontal wavenumber space. The wavenumbers are normalized by the deformation scale ($2\pi \times 16$ km).	43
Figure 2.13: Depth dependence of SQG-like KE spectra in the range of scales resolved by the ADCP data. The surface spectra follows a $\kappa^{-5/3}$ power law. The vertical dependence is computed from the SQG vertical structure for each κ given the global average stratification from the cDrake experiment. Colors represent spectra in different layers used in computing the ADCP and llc4320 spectral estimates.	44

Figure 3.1: (a) The study region with the subregion where the LLC outputs are analyzed. Colors represent the topography and white lines are contours of absolute dynamic topography every 0.1 m from AVISO. LLC 4320 (1/48°) snapshots of transects of potential density at 165°E (b and c) and surface vorticity (d and e). In (b) and (c) the white line represents the mixed layer depth. The snapshots were taken at 12:00 UTC.	48
Figure 3.2: Time series of the root-mean-square (RMS) of surface (a) vorticity, (b) rate of strain, and (c) horizontal divergence in the LLC outputs and gridded AVISO data. The convergence of meridians account for the small RMS divergence in AVISO line in (c).	51
Figure 3.3: Seasonal variation of joint probability distributions of surface lateral velocity gradient tensor: vorticity vs. strain rate (a through d), and vorticity vs. Laplacian of sea-surface height (e through h) in April (a, b, e, f) and October (c, d, g, h). Dashed lines in (a) through (d) represent one dimensional shear flow $\alpha = \pm\zeta$, characteristic of fronts. Dashed lines in (e) through (h) represent geostrophic flow $\zeta = \frac{g}{f} \nabla_h^2 \eta$	52
Figure 3.4: LLC4320 wavenumber-frequency spectrum of SSH variance in (a) April and (b) October. (c) Wavenumber spectrum of SSH variance — the integral of (a) and (b) over frequency. In (a) and (b), the light gray shaded region depicts the dispersion relations for inertia-gravity waves from mode 1 through mode 4 across the latitudinal domain; horizontal dashed lines represent the semi-diurnal lunar tidal frequency (M_2) and the inertial frequency at mid-latitude ($f_{32.5}$). For reference, the gray line shows the baseline requirement as specified by the SWOT science team [Rodriguez 2014].	54
Figure 4.1: Snapshots of the dipole solution with parameters in table 4.2. Contours depict potential vorticity, $q/(U_e k_e) = [-0.5, 0.5, 1.5]$, with dashed lines showing negative values. The upper row shows the wave action density $ \phi ^2/2f_0$. The lower row shows the wave buoyancy; the buoyancy scale is $B = k_e m U_w f_0 \lambda^2$. These plots only show the central $(1/5)^2$ of the simulation domain.	66
Figure 4.2: (a) Snapshot of q (contours) and ζ (colors) at $t \times U_e k_e = 20$. (b) Snapshot of q (contours) and wave PV q^w (colors). Both black lines and colors depict the contour levels $[-1.5, -0.5, 0.5, 1.5] \times (U_e k_e)$. Solid lines and reddish colors depict positive values; dashed lines and greenish colors show negative values.	69

Figure 4.3: Illustration of energy conversion terms in the dipole solution with parameters in table 4.2. (a) The action flux \mathcal{F} overlain on contours of relative vorticity $\Delta\psi/(U_e k_e) = [-1.5, -0.5, 0.5, 1.5]$, with dashed lines showing negative values; the scale of the action flux is $F = f_0 \lambda^2 k_e U_w^2$. (b) The local advective conversion (i.e., the unaveraged Γ_a) overlain on contours of streamfunction $\psi \times (U_e/k_e) = [-8, -4, -2, 0, 2, 4, 8]$.	73
Figure 4.4: Diagnostics of the Lamb-Chaplygin dipole solution with parameters presented in table 4.2. (a) Energy change about initial condition. (b) Wave potential energy budget (4.26).	74
Figure 4.5: Snapshots of the turbulence solution with parameters in table 4.4. Upper panels: PV q . Middle panels: wave kinetic energy density $ \phi ^2$. Bottom panels: wave buoyancy, with scale $B = k_e m U_w f_0 \lambda^2$. These plots show $(1/2)^2$ of the domain.	79
Figure 4.6: Diagnostics of the 2D turbulence solution with parameters presented in table 4.4. (a) Energy change about initial condition. (b) Wave potential energy budget (4.26).	80
Figure 4.7: Snapshots of PV q and its decomposition into relative vorticity $\zeta = \Delta\psi$, and wave potential vorticity q^w . The snapshots were taken at $t \times U_e k_e = 25$.	82
Figure 4.8: The energetics of 2D turbulence solutions with different dispersivities. (a) Energy change about the initial condition. (b) The energy conversion terms in (4.26). (c) The correlation between relative vorticity and wave kinetic energy. (d) The skewness of relative vorticity.	83
Figure 4.9: Energy-preserving spectra of 2D turbulence solutions with different dispersivities. The three panels show spectra of balanced kinetic energy \mathcal{K} , wave action \mathcal{A} , and wave potential energy \mathcal{P} . All solid lines correspond to spectra at $t \times U_e k_e = 25$ and the dashed line in \mathcal{K} is the balanced kinetic energy spectrum at $t \times U_e k_e = 0$. All spectra are normalized by their total energy, i.e., the area under each curve is one.	84
Figure 4.10: The escape of a Gaussian near-inertial wave packet from the saddle of a strain flow. The Gaussian decay scale is $\mu = \nu$. The waves are weakly dispersive: $\hbar = \eta/\alpha\mu^2 \approx 0.1$. Black contours show the streamfunction $\psi = -\alpha x y$, and colors represent wave zonal velocity; the colorbar limits are fixed. The black dot indicates the center of the packet and the gray line tracks its trail.	86

Figure 4.11: A comparison of passive-scalar and wave solutions ($\hbar \approx 0.09$) with same initial conditions (the wave kinetic energy is equal to the passive-scalar variance) and the same small-scale dissipation. (a) Initial condition of wave back-rotated zonal velocity and of the passive scalar. (b) Wave back-rotated zonal velocity at $t \times U_e k_e = 10$. (c) Passive-scalar concentration at $t \times U_e k_e = 10$. (d) Variance of wave velocity or passive-scalar variance. (e) Variance of wave velocity gradient or passive scalar. In (d) and (e), the diagnostics are normalized by their initial values.	87
Figure 4.12: Snapshot of the dipole Lagrangian-mean streamfunction and its decomposition into Eulerian-mean and Stokes drift. The snapshots is at at $t \times U_e k_e = 10$, just after the end of the refractive stage. The spatial average of each field was removed for direct comparison.	91
Figure 4.13: Diagnostics of the 2D turbulence solution with parameters presented in table 4.4. (a) Kinetic energy change about initial condition. (b) Refractive and advective conversion terms, and shear production.	92
Figure 4.14: Diagnostics of an illustrative solution with initially uniform wave and $\psi(t=0) = \sin x + \sin y$, with $\alpha = 0.1$ and $\hbar = 2$. (a) Energy difference about initial condition. (b) Wave potential energy budget.	96
Figure 5.1: Lagrangian-mean streamfunction ψ^L in (5.99), wave PV q^w in (5.97), and wave dissipation in (5.98) for Gaussian wave packets (5.92) with $ka = (0.1, 1, 10)$ and $l = 0$	119
Figure 5.2: Solution of the dissipative NIW-QG model (5.87)-(5.88) with initial conditions $q = 0$ and the Gaussian wave packet (5.92) with $ka = 1$ and $l = 0$	121
Figure 5.3: Energetics of a solution of the dissipative NIW-QG model (5.87)-(5.88) with initial conditions $q = 0$ and the Gaussian wave packet (5.92) with $ka = 1$ and $l = 0$. (a) Wave action, and (b) Lagrangian-mean and (c) Eulerian-mean kinetic energy budgets. The energy tendencies are normalized by $10^{-5} U_w^2 \gamma^{-1}$	125
Figure 5.4: (a) Spectrum of the random PV forcing, an annulus centered at k_f with width Δ_f . (b) A realization of the random PV forcing.	127
Figure 5.5: Snapshots at $t\gamma = 30$ of potential vorticity q and its decomposition into relative vorticity $\zeta \stackrel{\text{def}}{=} \Delta\psi^L$ and wave potential vorticity q^w defined in (5.81). (In the no-wave solution, $q^w = 0$ so that $q = \Delta\psi^L$.) The normalization constant is $Q = k_f U_e$ with $U_e = (\sigma_q^2/f_0)^{1/2}$. These plots show only $(1/2)^2$ of the simulation domain.	130

Figure 5.6: Time series of balanced kinetic energy \mathcal{K}^L , wave potential energy \mathcal{P} , and wave kinetic energy $f_0 \mathcal{A}$. \mathcal{K}^L and \mathcal{P} are normalized by the no-wave balanced kinetic energy prediction $E_q = \sigma_q^2/2\mu$, whereas $f_0 \mathcal{A}$ is normalized by the wave kinetic energy prediction $E_w = \sigma_w^2/2\gamma$, where $\gamma = \mu + m^2\nu$. Gray shading depicts the timespan used to compute averaged budgets.	132
Figure 5.7: Balanced kinetic energy budget of reference and no-drag solutions. The budget represents an average over $t\gamma \geq 10$, and each term is normalized by the stochastic forcing. The residual is less than 1% in both budgets.	133
Figure 6.1: Nondimensional base-state for the Eady problem using various truncation for the series (6.31). In the middle panel N is the number of baroclinic modes. (a) Zonal velocity: although the truncation has zero slope at the boundaries there are no Gibbs oscillations. (b) Meridional PV gradient associated with the truncated series (6.33). (c) as in (b) but with an expanded abscissa. As N increases, the PV gradient distributionally converges to two Brethertonian delta functions at the boundaries.	148
Figure 6.2: Growth rate for the Eady problem as a function of the zonal wavenumber ($l = 0$) using approximations A, B (exact), C with various number of baroclinic modes (N).	154
Figure 6.3: Structure of $\kappa = 8$ unstable mode for the Eady problem obtained using approximation A and $N = 64$. Streamfunction is the black curves and PV is the colors. The streamfunction slightly tilts westward as z increases. One can see the unphysical critical layer associated with the fast-oscillating base-state PV. The critical level, z_c , is the depth where the unstable wave speed matches the velocity of the base-state. Only the top quarter of the domain is shown.	158
Figure 6.4: Absolute error as a function of number of baroclinic modes (N) for the growth rates of the Eady problem. The solid lines show the error at the exact fastest growing mode ($\kappa \approx 1.6$). The dashed line is the approximation A error at $\kappa = 8$	159
Figure 6.5: Growth rate for the Green problem with $\hat{\beta} = 1$ as a function of the zonal wavenumber ($l = 0$) using approximations A, B, C with various number of baroclinic modes (N). The black line is a finite-differences solution with 1000 vertical levels.	160
Figure 6.6: Absolute error as a function of number of baroclinic modes (N) for the growth rates of the Green problem. The solid line represent the error at the exact fastest growing mode ($\kappa \approx 1.9$). The dashed line is the error at $\kappa = 8$	161

LIST OF TABLES

Table 2.1: Summary of data editing criteria applied to AltiKa data.	19
Table 2.2: Details of the llc spin-up hierarchy. In this study, we analyze a subset of the llc 4320 simulation is analyzed.	41
Table 3.1: The square root of the mean-square SSH at submesoscales (10-100 km) estimated from the wavenumber-frequency spectrum of SSH variance in figure 3.4. The SWOT baseline requirement at 15 km is approximately $\sqrt{2}$ cm.	55
Table 4.1: Summary of model-based studies of energy extraction from balanced flow by near-inertial waves.	59
Table 4.2: Description of parameters of Lamb-Chaplygin simulation. The initial condition have Rossby number $Ro = U_e k_e / f_0 \approx 0.05$, wave dispersivity $\hbar = f_0 \lambda^2 k_e / U_e \approx 1$, and wave amplitude $\alpha = Ro(U_w/U_e)^2 \approx 3.75$. .	67
Table 4.3: The time-integrated budget of wave potential energy and quasigeostrophic kinetic energy of the Lamb-Chaplygin dipole solutions with parameters provided in table 4.2. The energy budgets close within $10^{-6}\%$. .	75
Table 4.4: Description of parameters of the 2D turbulence simulations. The initial condition have Rossby number $Ro = U_e k_e / f_0 \approx 0.05$, wave dispersivity $\hbar = f_0 \lambda^2 k_e / U_e \approx 0.5 - 2$, and wave amplitude $\alpha = Ro(U_w/U_e)^2 \approx 0.1$	78
Table 4.5: The time-integrated budget of wave potential energy and QG kinetic energy of the reference 2D turbulence solution with parameters in 4.4. The energy budgets close within 0.1%.	78
Table 5.1: Parameters of the reference solution.	129

ACKNOWLEDGEMENTS

My greatest intellectual stimulation in graduate school came from my advisor, Bill Young, and unofficial co-advisors, Sarah Gille and Teri Chereskin. I thank Bill and Sarah and Teri for their sober advice, steadfast encouragement, and unabated enthusiasm over the years. I've drawn a lot of inspiration from their work ethic and academic achievement. And they deserve a lot of credit for my success.

Stefan Llewellyn-Smith and Juan C. Del Álamo joined the thesis committee upon my candidacy exam, and I appreciate their generous time, comments, and questions.

I'm also in debt with UCSD's congenial community. In particular, I thank my instructors at SIO and MAE for teaching me a great deal of oceanography and applied math; the 2013 COAP cohort, Greg Wagner, and Navid Constantinou for camaraderie (and coffee); and Tomomi Ushii, the SIO department office, the CASPO business office for administrative support.

My co-authors Ian Grooms, Dimitris Menemenlis, and Greg Wagner served as sounding boards during different stages of my thesis work. And chapters 2 and 3 would be incomplete without Dimitris's modeling efforts, with outstanding support from our colleagues at the NASA Advanced Supercomputing Division.

Finally, I thank my family and Nina for their loving support along the way.

My education and research at SIO were financed by the National Aeronautics and Space Administration, the National Science Foundation, and the Wyer Family Endowment.

Chapter 2, in full, reprints material as it appears in *Journal of Physical Oceanography*, 2016, 46, doi:10.1175/JPO-D-15-0087.1. Rocha, C. B.; Chereskin, T. K.; Gille, S. T.; Menemenlis, D. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, reprints material as it appears in *Geophysical Research Letter*, 2016, 43, doi: 10.1002/2016GL071349. Rocha, C. B.; Gille, S. T.; Chereskin, T. K.; Menemenlis, D. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in full, in full, reprints material as it appears in *Journal of Fluid Mechanics*, in press, doi:10.1017/jfm.2018.308. Rocha, C. B.; Wagner, G. L.; Young, W. R. The dissertation author was the primary investigator and author of this paper.

Chapter 6, in full, reprints material as it appears in Journal of Physical Oceanography, 2016, 46, doi:10.1175/JPO-D-15-0073.1. Rocha, C. B.; Young, W. R.; Grooms, I. The dissertation author was the primary investigator and author of this paper.

ABSTRACT OF THE DISSERTATION

The turbulent and wavy upper ocean: transition from geostrophic flows to internal waves and stimulated generation of near-inertial waves

by

Cesar Barbedo Rocha

Doctor of Philosophy in Oceanography

University of California, San Diego, 2018

Professor William R. Young, Chair

We study the mesoscale to submesoscale (10-300 km) dynamics of the upper ocean, with particular attention to the partitioning between geostrophic flows and internal waves, and the interaction between these two types of flow. Using 13 years of shipboard ADCP transects in Drake Passage, we show that internal waves account for more than half of the upper-ocean kinetic energy at scales between 10-40 km; a transition from the dominance of geostrophic flow to inertia-gravity waves occurs at 40 km. We further show that a global numerical model with embedded tides reproduces this partitioning between upper-ocean geostrophic flows and inertia-gravity waves. Using the output of this model, we show that in the Kuroshio Extension upper-ocean submesoscale (10-100 km) geostrophic flow and inertia-gravity waves undergo vigorous seasonal cycles that are out of phase: geostrophic flows peak in late winter/early spring, while the projection of inertia-gravity waves at the surface peaks in late summer/early fall.

The observational and modeling evidence of the importance of both geostrophic flows and internal gravity waves at mesoscales to submesoscales hints on the interaction between these two types of flow. To better understand these interactions, we analyze a simple model that couples barotropic quasi-geostrophic flow and near-inertial waves. There are two mechanisms of energy transfer from geostrophic flow to externally forced near-inertial waves: the refractive convergence of the wave action density into anti-cyclones (and divergence from cyclones); and the enhancement of wave-field gradients by geostrophic straining. Unforced inviscid numerical solutions of this reduced model

reveal that geostrophic straining accounts for most of stimulated generation, which represents 10-20% of the decay of the initial balanced energy. Consideration of the dissipative problem reveals that wave dissipation generates both quasi-geostrophic potential vorticity locally and geostrophic kinetic energy. And this wave streaming mechanism is non-negligible in forced-dissipative solutions, which equilibrate even without bottom drag.

In a separate study, we derive a Galerkin approximation for the surface-active quasi-geostrophic system using standard vertical modes. While the Galerkin expansions of streamfunction and potential vorticity do not satisfy the inversion relation exactly, the series converge with no Gibbs oscillations. With enough modes, the Galerkin series provide a good approximation to the streamfunction throughout the domain, which can be used to advect potential vorticity in the interior and buoyancy at the surfaces.

Chapter 1

Introduction

1.1 Background

The ocean absorbs, in the form of heat, about 90% of Earth’s radiation imbalance [e.g., Hansen et al. 2011]. And the ocean impact on Earth’s climate hinges on fine-scale processes that stir and mix heat and other tracers, thereby mediating the exchanges of these properties across the sea surface and between the upper ocean and the deep ocean. This dissertation investigates two of the most energetic of these processes: geostrophic eddies and internal waves.

Many quantitative oceanographic inferences of the 20th century were derived from moored instruments [Wunsch 2015]. Figure 1.1 depicts frequency spectra of currents observed by such an instrument attached to the WHOI699 mooring, which was deployed near the Kuroshio Current, near Japan. The spectra show that most of the ocean kinetic energy density lies on the subinertial band, with timescales longer than the inertial period, $2\pi/f_0$, where f_0 is the local rotation rate or inertial frequency. (The inertial period at mid latitudes is about one day.) Mesoscale eddies, with timescales of days to several weeks, dominate the ocean kinetic energy [e.g., Ferrari and Wunsch 2009]. In the WHOI699 mooring, in particular, 75% of the kinetic energy lies between periods of 2 days and 2 months, associated with mesoscale variability of the Kuroshio Current.

Mesoscale eddies are severely constrained by their aspect ratio, with vertical scales much smaller than horizontal scales, by Earth’s rotation, and by stable stratifica-

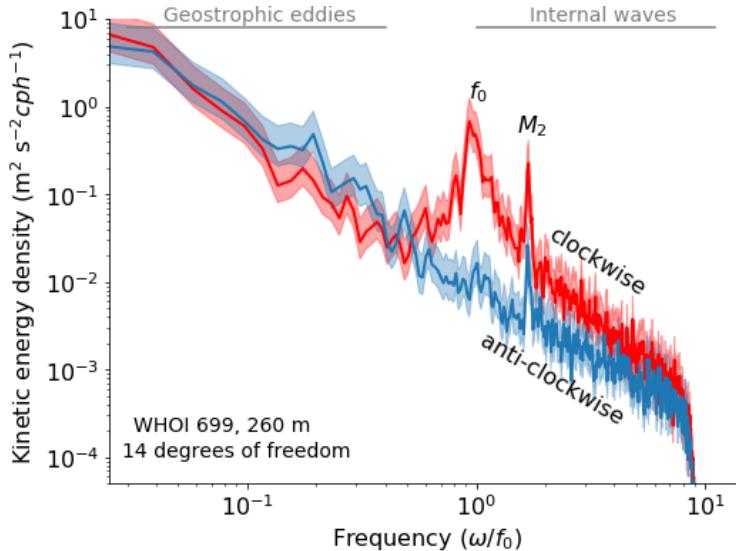


Figure 1.1: Rotary frequency spectra of kinetic energy from WHOI299 mooring at 260 m (nominal depth). The mooring was located at (34.98°N, 152.03°E), near the Kuroshio Current; the local depth was 6149 m. The red line shows the spectrum of clockwise motions and the blue line shows the spectrum of anti-clockwise motions. The inertial period is $2\pi/f_0 \approx 20.88$ h.

tion. These constraints render the motion approximately 2D, with horizontal currents near a state of balance between the Coriolis acceleration and the horizontal pressure gradient. At the sea surface this geostrophic balance is

$$(u_g, v_g) = \frac{g}{f_0} (-h_y, h_x), \quad (1.1)$$

where g is the acceleration due to gravity and h is the sea-surface height. Thus mesoscale surface currents can be *diagnosed* from sea-surface height observed by radar altimeters mounted on satellites. Indeed, satellite altimetry is responsible for a major observational leap in oceanography at the end of last century, revealing—on a global scale—the turbulent nature of the ocean circulation [e.g., Stammer 1997; Chelton et al. 2007]. Mesoscale eddies account for most of the isopycnal mixing in the ocean, and thanks to the geostrophic relation (1.1) the geography of lateral stirring and mixing has been mapped and its consequences for the large-scale circulation have begun to be investigated [e.g., Shuckburgh et al. 2009; Abernathey and Marshall 2013].

Another distinguishing characteristic of the spectra in figure 1.1 is the internal-wave band, with timescales sandwiched between the inertial period and the buoyancy

period $2\pi/N$, where N is the buoyancy frequency associated with the background stratification. (The WHOI699 mooring data were hourly averaged, so the buoyancy period, ~ 20 minutes, is not resolved.) Two distinct peaks account for most of the kinetic energy in the internal-wave band: near-inertial waves, with period near $2\pi/f_0 = 22.88$ h, and the M_2 tide, with period $2\pi/\omega_{M_2} = 12.42$ h. These frequencies stand out because they are directly forced by semi-diurnal barotropic lunar tides and by shifting winds at the sea surface [e.g., Ferrari and Wunsch 2009].

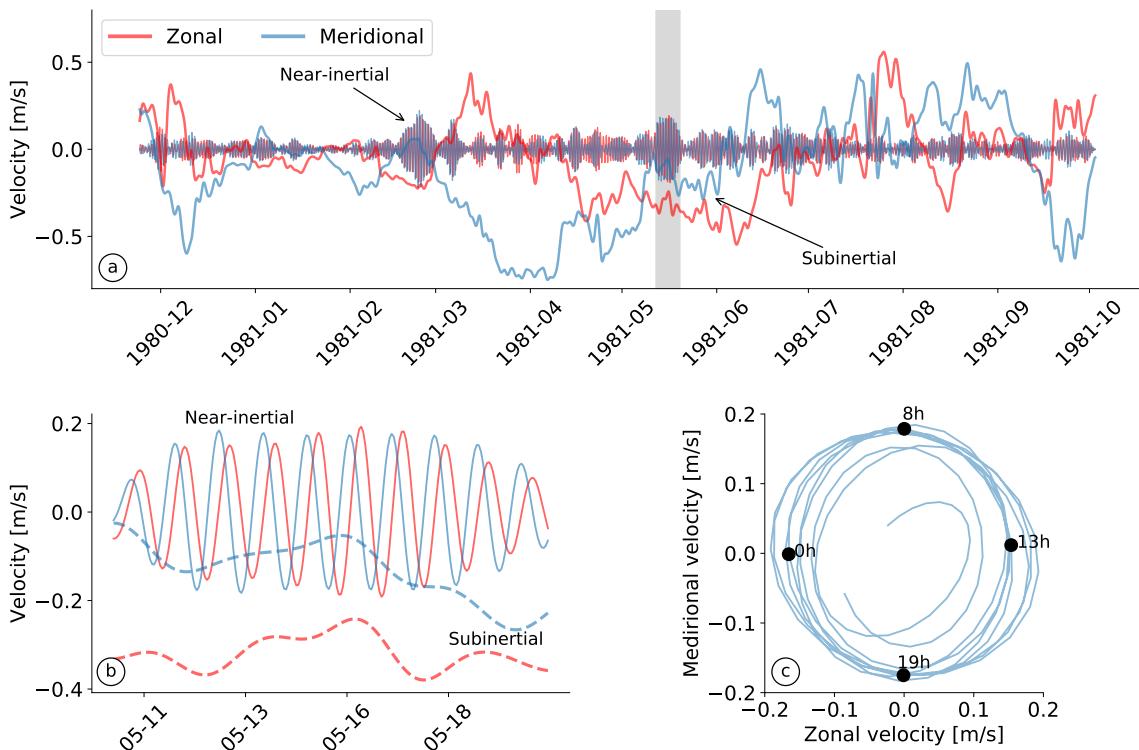


Figure 1.2: (a) Near-inertial and subinertial velocity time series from WHOI299 mooring at 260 m (nominal depth), whose spectra is shown in figure 1.1. (b) A blow-up of the shaded timespan in (a). (c) Near-inertial velocity hodograph for the velocity series in (b); black dots in (c) show data on May 15, 1981, depicting the clockwise polarization of the near-inertial currents.

In the WHOI699 mooring, the near-inertial peak, with frequencies $0.8f_0 \leq \omega_{ni} \leq 1.2f_0$, account for 14% of the total kinetic energy or about 50% of the eddy kinetic energy at frequencies higher than $0.8f_0$. The leaking of kinetic energy out of the inertial frequency towards lower frequencies probably stems from interactions of near-inertial waves with subinertial flows [e.g., Kunze 1985; Young and Ben-Jelloul 1997]. Near-

inertial motions are approximately circularly polarized and display an anti-cyclonic polarization, rotating clockwise in the northern hemisphere (see figure 1.1). These waves contain most of the ocean vertical shear and thus contribute significantly to diapycnal mixing [e.g., Munk 1981; Alford et al. 2016].

The WHOI699 mooring is not special: globally, mesoscale eddies account for 80-90% of the upper-ocean eddy kinetic energy, and internal waves contain most of the remaining energy, with near-inertial waves accounting for about half of the ocean internal-wave kinetic energy [e.g., Munk 1981; Ferrari and Wunsch 2009; Alford et al. 2016].

1.2 Motivation

While mesoscale eddies and internal waves are more easily separated by frequency, much less is known about how these flows project spatially, in particular on fine horizontal scales 10-200 km. This ignorance stems from lack of resolution in current satellite altimeters and paucity of in situ repeat spatial data [Fu and Ferrari 2008; Callies and Ferrari 2013]. Indeed, the upper-ocean horizontal wavenumber spectra at scales smaller than 100 km just began to be explored [Wang et al. 2010; Callies and Ferrari 2013].

A main motivation for studying the wavenumber projection of geostrophic flows and internal waves in the upper ocean is the upcoming Surface Water & Ocean Topography (SWOT) mission, with planned launch for 2021 [Fu and Ferrari 2008; Fu and Ubelmann 2014]. SWOT will employ a wide-swath synthetic aperture radar (SAR) technology to measure sea-surface height with 15-km resolution, up to 10 times finer than existing altimeters [Rodriguez 2014; Fu and Ubelmann 2014]. But it is unclear whether this increased horizontal resolution of sea-surface height will translate into an improved resolution of surface velocities: if internal-wave sea-surface signals rival those of geostrophic eddies, then diagnosing geostrophic velocities using (1.1) becomes moot. Understanding the transition from geostrophic flow to internal waves in the upper ocean will inform the planning of SWOT and help maximize the mission's scientific returns. Chapters 2 and 3 of this dissertation contribute to the SWOT pre-launch science

by assessing the relative contribution of upper-ocean internal waves and geostrophic flows on SWOT-relevant scales through analysis of observations and a numerical model output.

Another motivation for studying geostrophic flows and internal waves in the upper ocean is the unanswered question of how mesoscale eddies are dissipated. These eddies are energized by instabilities of wind-driven large-scale currents. And the energy supplied by the large-scale subinertial winds to the ocean circulation must be dissipated at the same rate by centimeter-scale viscous processes. But mesoscale eddies flux energy towards large spatial scales because they are approximately in geostrophic balance, so equilibration of the ocean circulation requires ageostrophic processes to transfer mesoscale energy towards small spatial scales; see McWilliams [2016] for a recent review. Several mechanisms that might result in this down-scale energy transfer have been extensively investigated, including spontaneous generation of internal waves by frontal instabilities, boundary layer turbulence, and lee-wave generation by mesoscale eddies impinging on rough topography [e.g., Nikurashin, Vallis, and Adcroft 2013; Nagai et al. 2015; McWilliams 2016]. But the mesoscale energy budget remains unclosed [Nagai et al. 2015].

A less studied mechanism is the direct transfer of energy from mesoscale eddies to existing internal waves [Gertz and Straub 2009]. Computational simulations and theory suggest that wind-driven near-inertial waves interact with geostrophic flows, with important energy exchanges [e.g., Gertz and Straub 2009; Xie and Vanneste 2015; Wagner and Young 2016]. This wave-mean interaction connects the low-frequency and high-frequency routes of windy energy in the ocean, potentially accounting for a major sink of mesoscale energy [Xie and Vanneste 2015]. Chapters 4 and 5 of this dissertation contribute to the physical understanding of this mechanism through theory and idealized computational simulations.

1.3 Dissertation's outline

This dissertation is written in manuscript style—chapters 2 through 6 are independent studies on geostrophic flows and internal waves in the upper ocean. Chapter

2 explores the mesoscale to submesoscale kinetic-energy spectrum in the Antarctic Circumpolar Current in Drake Passage. Using 13 years of near-surface velocity data from repeat shipboard ADCP transects, we show that the kinetic-energy spectrum is consistent with a combination of horizontally rotational and divergent flows. Several characteristics of the rotational component of the spectrum are consistent with interior quasigeostrophic turbulence. A $1/48^\circ$ global simulation with high-frequency forcing and embedded tides, which generate internal waves, reproduces the observed kinetic energy spectrum. Further analysis of this simulation reveals that inertia-gravity internal waves account for most of the horizontally divergent flows. The kinetic-energy spectrum thus transitions from geostrophic turbulence at mesoscales to internal waves at submesoscales, with a transition scale of about 40 km.

The observed and modeled kinetic-energy spectrum in Drake Passage shows little seasonal variability. But recent studies report strong upper-ocean variability in the subtropics. Chapter 3 thus presents a case study of upper-ocean seasonal variability in the Kuroshio Extension. Using surface velocity data from the $1/48^\circ$ global simulation, we show that the near-surface submesoscales turbulence and internal waves seasonal cycles are out of phase. Submesoscale turbulence peaks in later winter thanks to the deepening of the mixed layer through the stormy, cold season. And the near-surface variability due to internal waves peaks in summer owing to re-stratification.

The results of chapters 2 and 3 show that geostrophic flows and internal waves co-exist across mesoscales to submesoscales (10-200 km) in the upper ocean, thus hinting at the interaction of these two distinct types of flows. Chapters 4 and 5 use an asymptotic model that couples barotropic geostrophic flow with vertically planar near-inertial waves. Analysis of the coupled model shows that the reduction of lateral coherence of the waves is associated with a transfer of energy from the geostrophic flow to externally forced waves via two mechanisms: a “refractive conversion,” associated with the convergence of wave action into anti-cyclones; an “advective conversion,” which is a consequence of the lateral straining of the waves by the geostrophic flow. Chapter 4 details this stimulated generation of near-inertial waves using solution to initial-value problems in which the waves initially laterally uniform waves interact with decaying two-dimensional turbulence. The results reveal that the advective conversion

accounts for most of the energy transfers. But geostrophic straining is slow thanks to wave dispersion.

With the goal of studying the equilibration of 2D turbulence by stimulated generation, chapter 5 derives a dissipative vertical plane wave model starting from the dissipative Boussinesq equations. The dissipative effects on potential vorticity are non-monotonic, and wave dissipation can generate potential vorticity. With this dissipative model, we conduct a set of forced solutions in which the geostrophic flow is stochastically stirred and the waves are stochastically forced at large scales. Thanks to stimulated generation, 2D turbulence equilibrates even without bottom drag. Associated with the dissipative generation of potential vorticity, wave streaming contributes a non-negligible source of geostrophic kinetic energy.

Finally, in a separate study, chapter 6 presents the derivation of a consistent Galerkin approximation of the surface-active quasi-geostrophic equations using standard barotropic modes. While the Galerkin series of potential vorticity and streamfunction do not satisfy the inversion relation exactly, the series converge uniformly and absolutely. These Galerkin series can thus be used to correctly calculate the advection of potential vorticity and surface buoyancy. We illustrate this procedure using classical linear stability problems.

Chapter 2

Mesoscale to submesoscale wavenumber spectra in Drake Passage

2.1 Introduction

We investigate the governing dynamics at mesoscales to submesoscales in the upper ocean, with focus on the partition between balanced and ageostrophic flows. Specifically, we analyze upper ocean kinetic energy and sea surface height (SSH) variance spectra in the Antarctic Circumpolar Current (ACC) using 13 years of shipboard ADCP measurements in Drake Passage, complemented by satellite altimeter products and a 1-km numerical simulation with embedded tides.

This paper extends recent studies in the Gulf Stream [Wang et al. 2010; Callies and Ferrari 2013; Bühler, Callies, and Ferrari 2014] to the Southern Ocean, where differences in dynamical regimes and scales are expected. In particular, the ACC is organized in multiple fronts [e.g., Nowlin, Whitworth III, and Pillsbury 1977], and the combination of relatively high latitude and weak stratification results in a mean deformation radius of \sim 16 km and restricts the internal wave continuum to a range of periods from \sim 20 min to 14 h.

Following Callies and Ferrari [2013] (hereafter CF13), we employ two main diagnostics: spectral slopes of one-dimensional along-track wavenumber spectra and the ratio of across-track to along-track KE spectra. These diagnostics are applied to in situ velocity measurements (Section 2.3), synthetic data (Section 2.5) and a 1-

km primitive-equation numerical model (Section 2.6). Following Bühler, Callies, and Ferrari [2014], we split the one-dimensional KE spectra into rotational and divergent components and into inertia-gravity wave and geostrophic flow components (Sections 2.3 and 2.6). We further separate both synthetic and model velocity into horizontally divergent and rotational parts, and we compute KE and SSH variance isotropic spectra.

We find that the upper ocean KE spectra in Drake Passage are consistent with isotropic interior QG turbulence masked by horizontally divergent, ageostrophic flows. These ageostrophic motions are likely dominated by inertia-gravity waves and project onto horizontal scales as large as 150 km but are more significant at scales between 10 and 40 km where they account for about half of the KE. Model results indicate that horizontally divergent flows strongly imprint onto SSH.

2.2 Theoretical predictions

CF13 review the theoretical predictions for the oceanic horizontal wavenumber spectra at scales between 1 and 200 km. We summarize only the most fundamental results and emphasize the assumptions under which they are obtained.

The two turbulence paradigms invoked in the submesoscale range are interior quasigeostrophy (QG) and surface quasigeostrophy (SQG). Under the assumptions of statistical homogeneity and isotropy, QG turbulence theory predicts a κ^{-3} power law for scales smaller than the forcing scale, where $\kappa = \sqrt{k^2 + l^2}$ is the isotropic horizontal wavenumber [Charney 1971, CF13]. This forcing scale is generally associated with the wavelength of the most unstable mode of interior baroclinic instability [e.g., Salmon 1998]. Under the same assumptions, SQG turbulence predicts $\kappa^{-5/3}$ power law at scales smaller than the forcing scale [Blumen 1978, CF13]. It is not clear what would set the forcing scale of pure SQG turbulent flows in the ocean.

Shipboard transects and along-track altimeter observations permit estimation of only one-dimensional (along-track) wavenumber (k) spectra. In geostrophic turbulence, which is nondivergent to leading order, isotropic and one-dimensional wavenumber spectra follow the same power law (CF13). Furthermore, along-track (\hat{C}^ν) and across-

track (\hat{C}^u) velocity variance spectra are related through [Charney 1971]

$$\hat{C}^u = -k \frac{d\hat{C}^v}{dk}. \quad (2.1)$$

Thus, if the isotropic spectrum follows a κ^{-n} power law, then \hat{C}^v and \hat{C}^u are related through the scaling exponent n : $\hat{C}^u = n \hat{C}^v$. This important result was observed by Charney when comparing his predictions for geostrophic turbulence against data [Charney 1971]. The ratio $R \equiv \hat{C}^u / \hat{C}^v$ is thus a simple diagnostic that should be used together with spectral slopes to interpret one-dimensional KE spectra [CF13; Bühler, Callies, and Ferrari 2014].

Ageostrophic motions (inertia-gravity waves, mixed-layer instabilities, fronts, filaments, stratified turbulence, Langmuir cells, convection cells, etc) are also important in the submesoscale range. These flows typically flatten the spectra because they project onto small scales. A simple horizontally isotropic (empirical) model exists only for the spectrum of the internal wave continuum [the Garrett-Munk spectrum; Munk 1981]. Analysis of spectra in different vertical levels and isopycnals provides insight for interpreting wavenumber spectra dominated by different types of ageostrophic motions (CF13).

2.3 Velocity observations

2.3.1 The ARSV Laurence M. Gould ADCP data

We use data from the U.S. Antarctic Research and Supply Vessel (ARSV) Laurence M. Gould (LMG); the dataset is an update of the one used by Lenn et al. [2007] and by Lenn and Chereskin [2009]. The LMG crosses Drake Passage approximately twice per month on its way to Palmer Station, Antarctica, and back to Punta Arenas, Chile. While steaming at ~ 10 knots (~ 5 m s $^{-1}$), the LMG measures near-surface (26–300 m) currents using a hull-mounted 153.6-kHz ADCP. About 290 transects spanning 1999–2012 are available. The LMG research activities and weather conditions prohibit acquisition of data along an exactly-repeating transect; there are three main routes [Lenn et al. 2007]. The transects converge to about the same location in the north, the

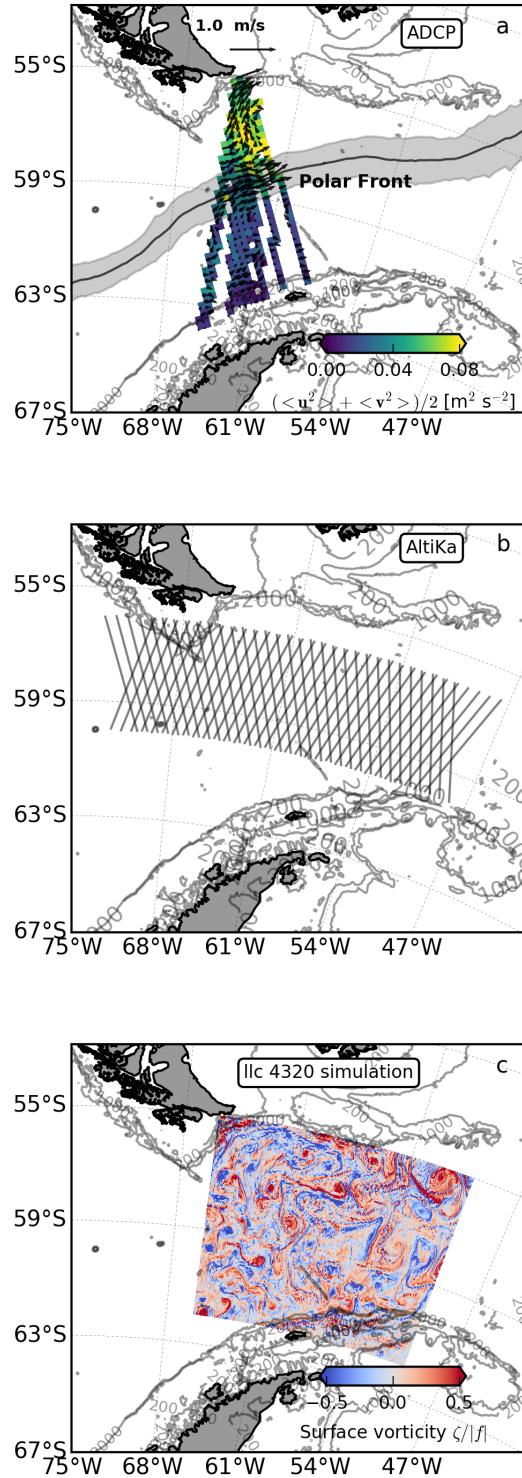


Figure 2.1: (a) Gridded mean (vectors) and eddy kinetic energy (colors) from ADCP data at 26 m (uppermost bin). The black line depicts the mean position of the Polar Front [1993-2013, updated from Sallée, Speer, and Morrow 2008]. (b) Segments of tracks from AltiKa altimeter. (c) Subdomain of llc4320 simulation used in this study. Colors represent a snapshot of surface vertical vorticity normalized by the local inertial frequency.

southern tip of South America, and spread out poleward (Fig. 2.1). The data are available at the “Joint Archive for Shipboard ADCP” (<http://ilikai.soest.hawaii.edu/sadcp/>).

The data were processed using the *Common Ocean Data Access System* Firing, Hummon, and Chereskin [2012, details at <http://currents.soest.hawaii.edu>]. In short, the ship-relative currents are rotated from the ship frame of reference to a geographic frame using the ship’s gyrocompass, corrected in post processing using Global Positioning System (GPS) attitude measurements. An accurate measurement of the ship’s velocity is obtained from high precision GPS. The speed of sound at the transducers, which are immersed in an anti-freeze solution, is determined from a sound velocity probe in the well. The final 5-min (~ 1.25 km), 8-m vertical bin average absolute ocean currents are accurate to ~ 2 cm s $^{-1}$ (appendix 2.A). We block-average the data every 5 km to make them evenly spaced, and this reduces the instrumental error to ~ 1 cm s $^{-1}$ (appendix 2.A).

The ship takes ~ 17 min to travel 5 km, ~ 5.5 h to cover 100 km, and ~ 2 days to cross the passage (~ 1000 km). We assume the ship velocity (~ 5 m s $^{-1}$) is much larger than the along-track propagation speed of the fluctuations [the fast-tow assumption, e.g., Wang et al. 2010, CF13]. For buoyancy frequencies of 3.5×10^{-3} rad s $^{-1}$ and 6.0×10^{-3} rad s $^{-1}$ [Argo climatology; dataset produced by Roemmich and Gilson 2009], the minimum period cut-off for free internal gravity waves (IW) is ~ 30 min to the north and ~ 17 min to the south of the Polar Front. The fast-tow assumption is reasonable for high-frequency internal waves because these flows project onto small scales (< 50 km). Linear first mode internal tides have a propagation speed of ~ 1.5 m s $^{-1}$ (see section 2.7), and are likely aliased. We assess the effects of aliasing of high-frequency motions on spectral slopes using data from a numerical model (Section 2.6). Moreover, removing barotropic tides using estimates from a shallow water model constrained to satellite altimeter data [TPXO7.2, updated from Egbert, Bennett, and Foreman 1994] does not change the ADCP spectral characteristics in Drake Passage; the M2 constituent dominates the tidal signal, with currents at the instrumental noise level (< 2 cm s $^{-1}$). The spectra presented here include barotropic tides.

We select 800-km-length, gap-free transects; no interpolation is performed.

For each transect, at each depth, the velocity components are rotated onto an along-track/across-track coordinate system. We then remove the spatial mean and linear trend, multiply each transect by a Hanning “spectral window”, and compute the discrete Fourier transform; results are insensitive to the choice of window. We repeated the calculations mirroring each transect to make it periodic. Spectra of “periodized” transects are indistinguishable from the pre-windowed spectra within error bars. The spectra is corrected for the variance removed by the window. We do not remove the low-resolution gridded mean; removing the mean changes the spectra only at scales larger than ~ 200 km.

We form spectral estimates by multiplying the Fourier coefficients by their complex conjugate and averaging over all realizations in three different layers (26-50 m, 58-98 m, and 106-202 m). Climatologically, the mean mixed-layer depth (MLD) in Drake Passage is $\sim 40 \pm 10$ m in summer and $\sim 120 \pm 30$ m in winter, with deeper MLDs north of the Polar Front [Argo mixed layers; dataset produced by Holte et al. 2010; Holte and Talley 2009]. Hence the deepest layer considered in this study is deeper than the mixed-layer (ML) except in some realizations in winter. The weak stratification in Drake Passage, however, makes MLDs difficult to determine; different methods lead to inconsistent MLDs, particularly north of the Polar Front [Stephenson, Gille, and Sprintall 2012]. Because of this caveat, we avoid discussing the results in terms of ML/pycnocline motions; we explicitly state the range of depths in each layer. Finally, Ekman currents are relatively small in Drake Passage (< 2 cm s $^{-1}$) and decay significantly in the upper 50 m [Lenn and Chereskin 2009], and therefore are likely present only in the 26-50 m layer.

Given the length of transects and resolution of the reduced data, the ADCP KE spectra resolve horizontal variability at scales spanning 1.0×10^{-1} cycles per km, cpkm, (10 km; Nyquist scale) to 1.25×10^{-3} cpkm (800 km; fundamental scale) with spectral resolution of 1.25×10^{-3} cpkm. Formal errors are obtained assuming that the Fourier coefficients are normally-distributed (details in appendix 2.B).

We employ a recently proposed Helmholtz decomposition to split the spectra into horizontally rotational and divergent components [Bühler, Callies, and Ferrari 2014]. This decomposition assumes that the flow is statistically stationary, homoge-

neous, and horizontally isotropic, and the fields are doubly periodic [Bühler, Callies, and Ferrari 2014]. Under these assumptions, the unique decomposition is achieved given the across-track $\hat{C}^u(k)$ and along-track $\hat{C}^v(k)$ KE spectra estimated from observations (details in appendix 2.C). Following Bühler, Callies, and Ferrari [2014], we further separate the rotational and divergent spectra into inertia-gravity wave and geostrophic flow components. Since no buoyancy data of sufficient quality and horizontal resolution concurrent with the ADCP data are available, we use the Garrett Munk spectrum for inertia-gravity waves [e.g., Munk 1981] to achieve the second step of the Bühler, Callies, and Ferrari [2014] decomposition (appendix 2.C).

2.3.2 Results

We present along-track wavenumber spectral estimates for along-track and across-track velocity components (Fig. 2.2). (In all figures, KE spectra are plotted, i.e., half of the across-track and along-track velocity variance spectra.) The spectra are steep at scales smaller than 200 km and flatten out at larger scales. Although this flattening at large scales could represent an actual change in inertial range or in the dynamics, the finiteness of the transects may account for part of this flattening. We only discuss results at scales between 10 and 200 km.

The across-track KE spectrum ($\hat{C}^u/2$, the red line in Fig. 2.2) is just slightly flatter than k^{-3} at scales smaller than 200 km. The along-track KE spectrum ($\hat{C}^v/2$, the blue line in Fig. 2.2) is flatter than \hat{C}^u but steeper than k^{-2} . The across-track KE spectrum \hat{C}^u is larger than the along-track KE spectrum \hat{C}^v at scales between 40 and 200 km. The two spectra \hat{C}^u and \hat{C}^v are indistinguishable below 40 km. The across-track to along-track KE ratio R varies from ~ 2.7 at 170 km to ~ 1.0 at 40 km. The average value of R at scales between 40 and 200 km is 1.52 ± 0.06 , 1.52 ± 0.06 , and 1.54 ± 0.08 for depth intervals of 26-50 m, 58-98 m, and 106-202 m, respectively.

The Helmholtz decomposition of Bühler, Callies, and Ferrari [2014] indicates an equipartition between rotational (\hat{C}^ψ , the green line in Fig. 2.2) and divergent (\hat{C}^ϕ , the yellow line in Fig. 2.2) KE components at scales between 10 and 40 km. The rotational component \hat{C}^ψ dominates at scales larger than 40 km. The divergent component \hat{C}^ϕ accounts for only $\sim 10\%$ of the total KE at 200 km.

Both KE components \hat{C}^u and \hat{C}^v are more energetic across all scales in the 26–50 m depth range compared to deeper levels. However, the spectral shape does not present any obvious depth dependence, i.e., spectral slopes and the across-track to along-track ratio R are very similar across all layers. KE spectra from a 38-kHz ADCP recently installed on the LMG [Firing, Hummon, and Chereskin 2012] suggest that these spectral characteristics extend to depths of \sim 1000 m; uncertainties are much higher owing to a small number of transects, and therefore we do not discuss these results here.

To assess the effect of the Polar Front on the spectra (e.g., rendering the results statistically inhomogeneous) we computed spectra using 250-km-length sub-transects north and south of the climatological Polar Front (PF, outside the gray shaded region in Fig. 2.1). These spectra resolve variability at scales spanning 10–250 km, with resolution of 4×10^{-3} cpkm. While the region north of the PF is more energetic, as also seen in the physical domain (Fig. 2.1), the spectral shape and the across-track to along-track R are very similar in both sub-regions (Fig. 2.3). The collapse of the KE components \hat{C}^u and \hat{C}^v into a single curve seems to extend to larger scales south of the PF. More years of data are required to assess this potential difference with statistical significance.

The wave-vortex decomposition assuming the Garrett-Munk spectrum for the inertia-gravity wave component yields a consistent picture (Figure 2.2d-2.2f). In particular, the geostrophic flow component dominates at scales larger than 100 km, whereas the inertia-gravity wave component dominates at scales smaller than 40 km; at scales between 40 and 100 km there is an approximate equipartition between vortex and inertia-gravity wave kinetic energy (Figure 2.2d). The spectra of the geostrophic flow component are consistent with predictions of isotropic interior quasigeostrophic turbulence from scales between 10 and 200 km, with steep spectra approximately following a k^{-3} power law, and the across-track component about three times larger than the along-track component (Figure 2.2e).

The ADCP data show no seasonal variability. At scales smaller than 40 km, the KE spectra seems shallower in austral summer than in austral winter (not shown). But this result is insignificant at the 95% confidence level; there are only about 30 DOF

in the seasonal spectra; gaps are more ubiquitous in austral winter owing to high sea conditions. This apparent lack of seasonality contrasts with the Gulf Stream region, where injection of submesoscale energy significantly flattens the KE spectrum in winter [Callies et al. 2015]. An accurate assessment of seasonality in the Drake Passage spectra would require more data.

In summary, one-dimensional KE spectra in the ACC in Drake Passage are reminiscent of predictions of isotropic interior QG turbulence. Both KE spectra \hat{C}^u and \hat{C}^v approximately follow a k^{-3} power law between 10 and 200 km. However, $R \approx 1.5 < 3$, suggesting that ageostrophic flows may be significant across this range of scales. The ageostrophic motions are likely dominated by inertia-gravity waves, and are dominant at scales smaller than 40 km. We further explore this question with the aid of stochastically-generated velocity fields and a numerical model (Section 2.6).

2.4 Satellite altimetry

2.4.1 The SARAL/AltiKa altimeter

Altimeter data for this analysis are drawn from the SARAL/AltiKa altimeter, which is a French and Indian mission launched in February 2013 Verron et al. [2015]. The altimeter data were produced and distributed by AVISO (<http://www.aviso.altimetry.fr/>), as part of the Ssalto ground processing segment. AltiKa was selected for this study, because its Ka-band altimeter has a footprint of about 8 km, compared with 20 km footprints in the Ku-band and C-band altimeters used for the TOPEX/Poseidon/Jason and ERS/Envisat series. In addition, in contrast with CryoSat-2, AltiKa flies on an exact repeat orbit, allowing us to remove the time mean sea surface height and to compute wavenumber spectra from sea surface height anomalies. Verron et al. [2015] show that wavenumber spectra from AltiKa agree well with Jason-2 and Cryosat wavenumber spectra for lengthscales longer than about 70 km. AltiKa reports sea surface height data at 40 Hz, equivalent to about 40 data postings every 7 km, a spatial sampling rate of about 0.177 km. The AltiKa spectra for the Drake Passage region are computed from 18 ascending and 18 descending \sim 460-km-length track segments (Figure 2.1c). Thus altimeter spectra presented here represent scales from

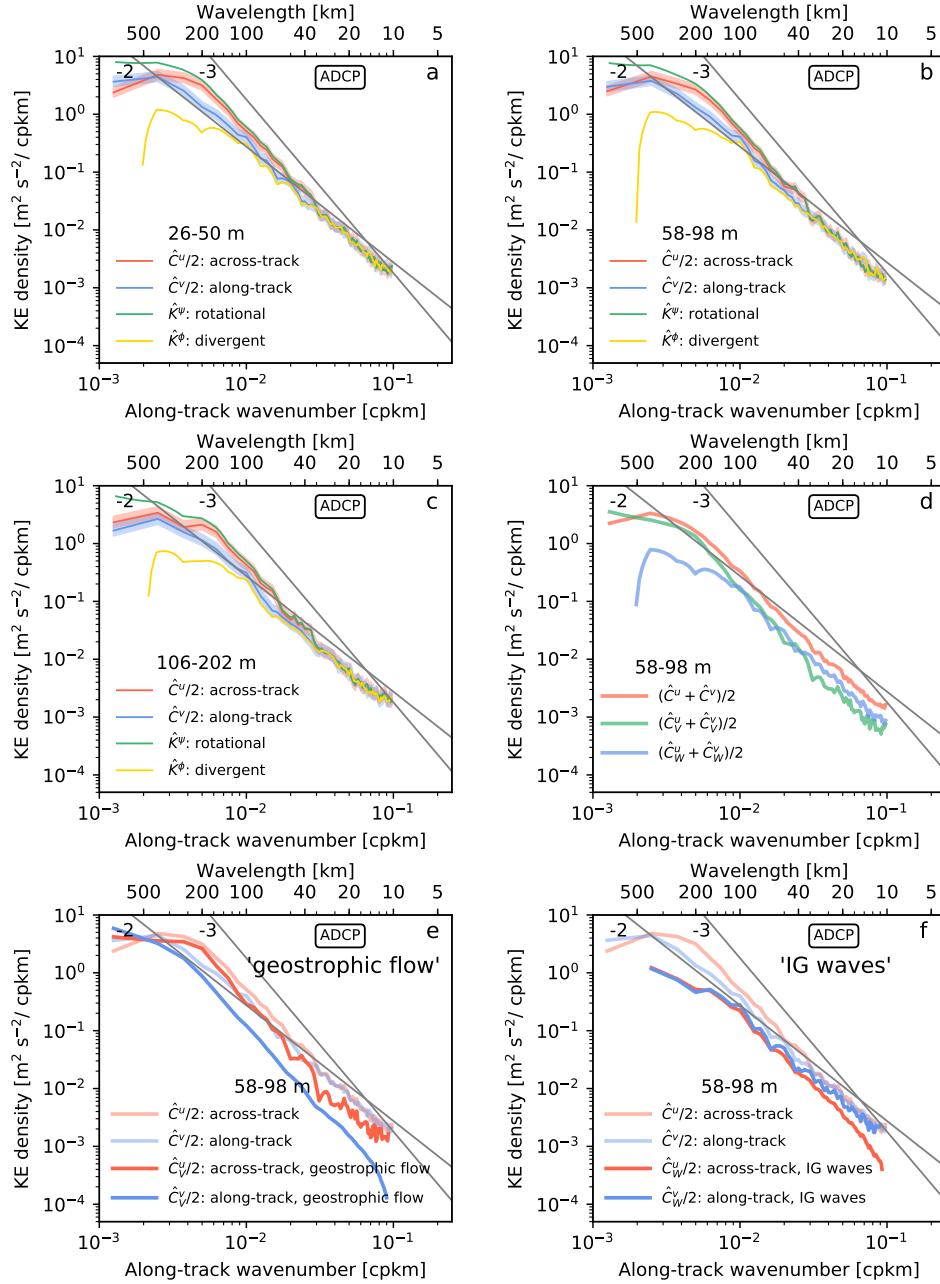


Figure 2.2: ADCP KE spectra for three layers: (a) 26-50 m, (b) 58-98 m, and (c) 106-202 m; and wave-vortex decomposition for the 58-98 m layer: (d) total kinetic energy, (e) the geostrophic flow (vortex) and (f) the inertia-gravity wave components of across-track and along-track components. The figure depicts across-track (red line) and along-track (blue line) KE spectra, and the spectral decomposition into horizontally rotational (green line) and divergent components (yellow line). Shaded represent 95% confidence limits. For reference, k^{-2} and k^{-3} curves are plotted (gray lines).

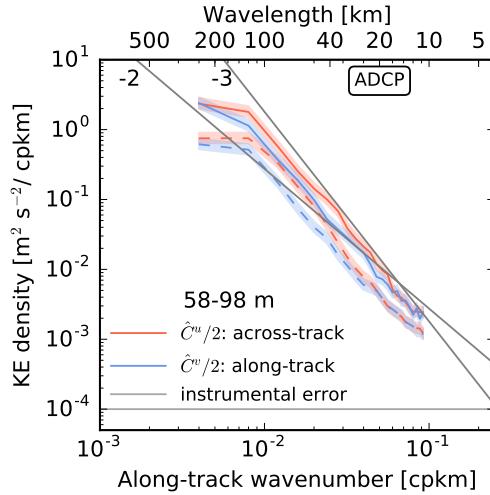


Figure 2.3: ADCP KE spectra for the 58-98 m layer from sub-transects to the north (solid) and to the south (dashed) of the Polar Front.

0.022 cpkm (460 km) to about 2.9 cpkm (~ 0.35 km).

The data are processed using an approach similar to that applied to coastal oceanography to edit data and retain small-scale signals Durand et al. [e.g., 2008], Roblou et al. [2011], and Birol, Cancet, and Estournel [2010]. Data are flagged as questionable if they are outside the parameter ranges specified in table 2.1 or if the data quality flags are anything other than “good” (with the exception of European Centre for Medium-range Weather Forecasting, ECMWF, data availability, for which we require a flag of 2 or less.)

We use the first 20 cycles of AltiKa data, which extend from 14 March 2013 through 11 February 2015. Wavenumber spectra are computed analogously to the ADCP spectra (section 2.3).

2.4.2 Results

Figure 2.4 shows the AltiKa along-track wavenumber spectra of SSH variance. Spectra of ascending and descending tracks are indistinguishable within error bars. The SSH variance spectra are steeper than k^{-4} at scales between 60 and 200 km. For scales smaller than 60 km, AltiKa wavenumber spectra have a stepped structure that

Table 2.1: Summary of data editing criteria applied to AltiKa data.

Surface type (0=ocean or semi-enclosed sea)	= 0
Number of valid points for range (count)	≥ 20
Root-mean-squared range (m)	0 to 0.2
Sea surface height anomaly (m)	≤ 2
Dry troposphere correction (m)	-2.5 to -1.9
Wet troposphere correction (m)	-0.5 to -0.02
Ionosphere correction (m)	-0.1 to 0.04
Sea state bias (m)	-0.5 to 0.0025
Ocean tide (m)	≤ 5
Solid earth tide (m)	≤ 1
Pole tide (m)	≤ 0.015
Significant wave height (m)	0 to 11
Backscatter coefficient σ_0 (dB)	3 to 30
Wind speed ($m s^{-1}$)	0 to 30
Off-nadir angle squared (degrees 2)	-0.2 to 0.015
rms(σ_0) (dB)	≤ 1
Number of valid points to compute σ_0 (count)	≥ 20

is characteristic of nearly all altimeters [e.g., Faugere et al. 2006], with the exception of Cryosat, which has a narrow footprint. AltiKa spectra are white for scales between about 15 and 5 km and then drop off in energy before plateauing again to be white at scales smaller than 2-3 km. For scales smaller than 2-3 km, the low-energy white spectra represent instrument noise stemming from the fact that the AltiKa footprint is larger than the spacing between measurements, so consecutive measurements are effectively repeated looks at the same patch of ocean. The intermediate white spectrum for scales between about 5 and 15 km, is referred to as a spectral hump and is associated with blooms of normalized radar backscatter (σ_0) or rain contamination Dibarboore et al. [2014]. While there have been suggestions for strategies to remove the spectral hump [e.g., Xu and Fu 2012; Zhou, Wang, and Chen 2015] to obtain consistent spectral slopes, we opt not to apply these approaches here, and instead truncate our spectral interpretation at about 30 km.

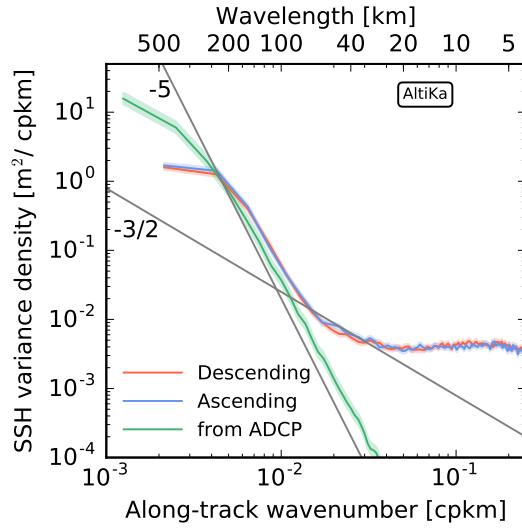


Figure 2.4: AltiKa SSH variance along-track wavenumber spectra. For comparison, also shown is the equivalent spectrum inferred from the ADCP spectrum by assuming geostrophy.

2.5 Some experiments with synthetic data

2.5.1 Synthetic velocity fields

We use stochastically-generated velocity fields to gain insight into the interpretation of the ADCP spectra and to test the accuracy of the spectral Helmholtz decomposition (Section 2.3). We create two-dimensional model spectra such that the associated isotropic spectra follow κ^{-3} and κ^{-2} power laws. These spectra are combined with Monte Carlo-generated random phases to create synthetic velocity fields [similar to the procedure employed by Rudnick and Davis 2003]. We employ a Helmholtz decomposition to separate the two-dimensional vector fields into rotational and divergent components [e.g., Capet et al. 2008]. Because we generate the synthetic data on a two-dimensional regular grid, the Helmholtz decomposition is efficiently performed in Fourier space [e.g., Smith 2008]. We calculate one-dimensional KE spectra as in section 2.3. We average spectra of 1000 independent velocity fields so that formal errors are insignificant.

2.5.2 Results

As expected for a purely horizontally rotational velocity field with a κ^{-3} isotropic spectrum, the across-track KE spectrum \hat{C}^u is three times larger than the along-track KE spectrum \hat{C}^v , i.e., $R = 3$ (Figure 2.5a). Adding a high-pass (40-km cut-off) component of horizontally divergent velocity field with a κ^{-3} isotropic spectrum produces KE spectra that resemble those from the ADCP (solid yellow and green lines in Figure 2.5b). The presence of divergent flows at high wavenumbers with the same level of variance as rotational flows results in $R = 1$ at scales smaller than 40 km. At larger scales, the rotational component dominates.

No doubt, other combinations produce similar spectra. For instance, adding the velocity field with a κ^{-3} isotropic spectrum to high-pass (40 km) divergent velocity field with a κ^{-2} isotropic spectrum and a variance that is 25% of that of the rotational flow generates very similar spectra (dashed yellow and green lines in Figure 2.5c). However, in agreement with the Helmholtz decomposition of the ADCP spectra, this simple red noise analysis suggests that in Drake Passage, divergent motions alter the QG turbulence spectral characteristics. The analysis of the numerical model supports this interpretation (Section 2.6).

The Helmholtz decomposition of the synthetic data spectra produces consistent separation into rotational and divergent components (Figure 2.5). The small ($\sim 10\%$) residual at scales larger than 40 km stems from the filtering process; residuals for purely rotational spectra are much smaller (Figure 2.5b and c). Thus this experiment with synthetic data verifies the robustness of the Bühler, Callies, and Ferrari [2014] decomposition provided that the fields have stationary, homogeneous, and horizontally isotropic statistics.

2.6 A high-resolution primitive-equation model

2.6.1 The llc4320 MITgcm simulation

We analyze model output from a primitive-equation (PE) global ocean simulation that resolves mesoscale eddies, internal tides, and other hydrostatic processes at

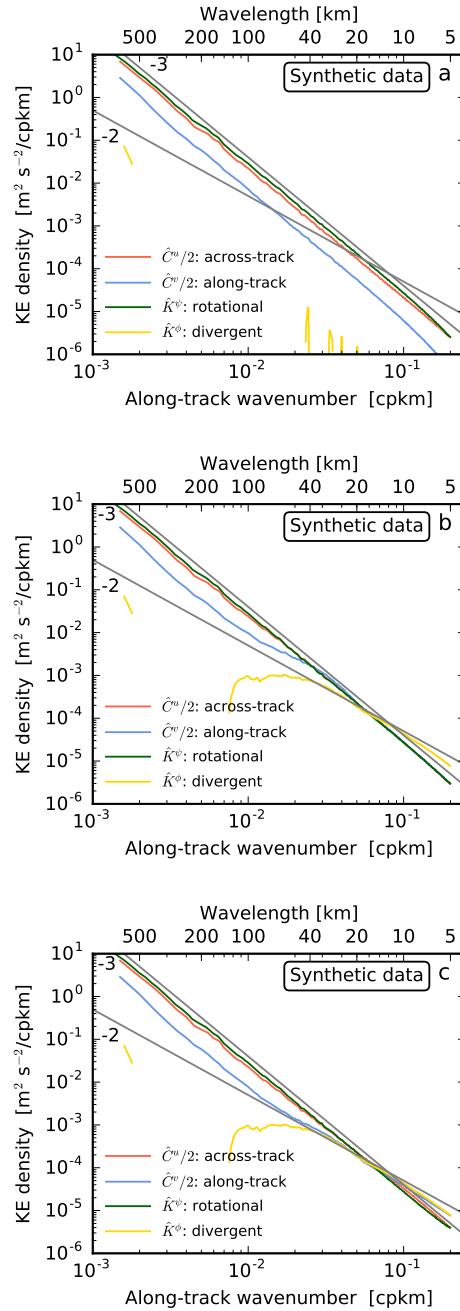


Figure 2.5: Synthetic data KE spectra and its decompositions into rotational and divergent components for different combinations of horizontally rotational and divergent flows. (a) κ^{-3} rotational flow, (b) κ^{-3} rotational flow combined with a high-pass (> 40 km) κ^{-3} divergent flow, and (c) κ^{-3} rotational flow combined with a high-pass (> 40 km) κ^{-2} divergent flow. In (c), the energy level of the divergent flow is a quarter of the energy level of the rotational flow. For reference, k^{-2} and k^{-3} curves are plotted (gray lines).

scales as small as about 5 km. The llc4320 is the highest resolution ($1/48^\circ$) of a hierarchy of simulations. We use the 84 days of llc4320 model output spanning September 2011 through December 2011 that were available at the start of this study. The model is forced by high-frequency winds and other surface fluxes, and tides (details in appendix 2.D). The average horizontal grid spacing in Drake Passage is ~ 1.1 km.

We emphasize the novelty of the global simulation analyzed here. In particular, the llc4320 simulation fully resolves mesoscale eddies and internal tides and admits some smaller scale variability. This contrasts with the simulation analyzed by Sasaki et al. [2014], which has coarser resolution and neglects tidal forcing altogether. Thus while the Sasaki et al. [2014] simulations represent mainly balanced dynamics, the llc4320 is able to support a more realistic mix of balanced and ageostrophic motions.

We perform our analysis in a sub-domain spanning $65\text{--}50^\circ\text{W}$, $55\text{--}62^\circ\text{S}$ (Fig. 2.1); we chose the eastern and polar limits to avoid sea-ice contamination. Before computations, we linearly interpolate the fields onto a regular Cartesian grid (1.15×1.15 km). We compute one-dimensional wavenumber spectra similarly to the ADCP, altimeter, and synthetic data estimates (sections 2.3–2.5). The spectral resolution is similar to the ADCP (Section 2.3). The Nyquist wavenumber is ~ 0.5 cpkm but we restrict our discussion to scales larger than the model viscosity cut-off of ~ 5 km (see dramatic change in spectral slopes in Fig. 2.7).

We also estimate horizontal isotropic wavenumber spectra from the model fields. We remove the time-mean and spatial averages of the horizontal fields. We then compute the 2D discrete Fourier transform (DFT) of these fields multiplied by a 2D Hanning window. We form the spectra by multiplying the Fourier coefficients by their complex conjugates, and averaging over all realizations. We form 2D kinetic energy spectra by summing spectral estimates of the two velocity components and dividing by two. The 2D spectra are then interpolated from Cartesian to polar coordinates (in wavenumber space) and integrated in the azimuthal direction [e.g., Richman et al. 2012]; in performing this interpolation, we impose low and high wavelength cut-offs of ~ 5 and ~ 500 km, respectively.

2.6.2 Results

There are clear inconsistencies between model¹ and ADCP KE spectra (Fig. 2.6). The model is more than twice as energetic as the ADCP data, and, among other differences, the model spectra peaks at about ~ 150 km, whereas the ADCP \hat{C}^v peaks at ~ 250 km. Besides being based on observations, statistics from the ADCP are more reliable because they are based on realizations taken over 13 years as opposed to 84 days of model simulation. There are too few independent realizations in 84 days of simulation to produce robust spectra at low wavenumbers (see appendix 2.B). Model KE spectra should be interpreted with these caveats.

First, we assess the degree to which the fast-tow assumption and the block-averaging of the ADCP data affect the spectra. We re-order the model output so that the end of a transect represents data sampled 2 days after its beginning. To mimic the ADCP measurements, we sample transects starting half from the north and half from the south. At scales smaller than 150 km, the spectrum of this “LMG sampled” model velocity (solid green line in Fig. 2.7) is indistinguishable from the spectrum of the instantaneous field (red line in Fig. 2.7). Similar results hold for the along-track spectrum (not shown). For instance, spectral slopes at scales between 40 and 150 km change from -2.51 ± 0.12 to -2.52 ± 0.12 for \hat{C}^u , and from -2.51 ± 0.12 to -2.36 ± 0.12 for \hat{C}^v ; R changes from 1.99 ± 0.05 to 1.87 ± 0.05 . Block-averaging the sampled fields decreases the KE at scales smaller than 20 km, slightly steepening the spectrum at high wavenumbers (dashed yellow and green lines in Fig. 2.7). Thus, aliasing of rapidly evolving flows due to the ship’s finite velocity does not change significantly the spectral slopes.

Model KE spectra lack a systematic depth dependence in the upper 200 m (Fig. 2.8), consistent with the ADCP KE spectra (Fig. 2.2). The model spectra are, however, flatter than the ADCP KE spectra at scales between 10 and 200 km; the model spectral slopes are flatter than k^{-3} but steeper than k^{-2} . Also consistent with the ADCP spectra, the model along-track spectra \hat{C}^v are flatter than the across-track spectra \hat{C}^u , and the average ratio of across-track to along-track spectra, R , is ~ 1.9 at scales between

¹KE spectra for zonal wavenumber spectra have very similar characteristics, with meridional (across-track) KE typically larger than zonal (along-track) KE.

40 and 150 km. At small scales, however, \hat{C}^u and \hat{C}^v do not collapse into a single curve. Instead, the along-track component \hat{C}^v dominates at scales smaller than 25 km, consistent with the dominance of the divergent component \hat{C}^ϕ (yellow lines in Fig. 2.8). Conversely, the rotational spectrum \hat{C}^ψ dominates at scales larger than 40 km. The local peak in divergent KE is likely associated with low-mode internal tides (see discussion in Section 2.7).

Removing high-frequency motions significantly changes the model KE spectral characteristics and the partition into rotational and divergent components. KE spectra of daily-averaged fields (dashed lines in Fig. 2.8) are much steeper than the spectra of hourly fields (solid lines in Fig. 2.8). For instance, in the 106-202 m layer, at scales between 40 and 150 km, the spectral slope changes from -2.58 ± 0.12 to -2.87 ± 0.12 for the across-track component \hat{C}^u . Moreover, the ratio of across-track to along-track R increases from 2.01 ± 0.05 for hourly velocity to 3.12 ± 0.10 daily-averaged fields. The ageostrophic motions are efficiently removed by daily-averaging the fields. The rotational component \hat{C}^ψ dominates at scales larger than 10 km. The spectra of daily-averaged velocity in the 10-150 km range have characteristics consistent with predictions of isotropic interior QG turbulence (e.g. CF13).

Conclusions based on analyses of the model isotropic spectra are consistent with results from the one-dimensional model spectra. The KE spectra of hourly fields follow a power law between κ^{-2} and κ^{-3} (Fig. 2.8d-2.8f). Removing high-frequency motions suppresses high-wavenumber variability; the spectra of daily-averaged fields are steeper than κ^{-3} , reminiscent of QG turbulence (e.g. CF13).

A Helmholtz decomposition of the two-dimensional velocity fields shows that a significant fraction of the spectral flattening at high wavenumbers stems from horizontally divergent, ageostrophic motions. The rotational component of the KE spectrum approximately follows a κ^{-3} power law at scales between 10 and 150 km. At the surface, the spectrum of the divergent flows peaks at about 150 km. These ageostrophic flows account for about 20% of the KE at scales between 10 and 150 km, and they contain about half of the KE at scales smaller than 40 km. Divergent flows dominate the KE at scales smaller than 20 km. High-frequency motions account for a significant fraction of divergent flows. This suggests that the divergent component is dominated by

inertia-gravity waves. The KE spectra of daily-averaged fields are consistent with those of rotational fields. Moreover, the residual isotropic KE spectrum (hourly minus daily averaged) has a shape consistent with the Garrett-Munk spectrum at scales between 10 and 100 km (Figure 2.8f).

To further assess the interpretation that most of the ageostrophic high wavenumber flows are accounted for by inertia-gravity waves, we calculate a horizontal wavenumber-frequency spectrum. Focusing on the high frequency content, we estimate the spectrum every 10 days, and average the results to form the spectral estimate. Because the spectrum is horizontally isotropic, we also integrate over azimuthal direction. The results show that at scales larger than 100 km there is a gap between the energetic low frequency (sub tidal) flows and the high frequency flows (Figure 2.10); no clear gap exists at smaller scales. At scales smaller than 40 km, 49% of the KE is accounted for by flows with frequencies larger than 0.9 f. Most of this high frequency energy is spread about the dispersion relationship of linear inertia-gravity waves (Figure 2.10).

The KE spectrum computed from velocities diagnosed from SSH is consistent with surface velocity fields at scales larger than 40 km (Fig. 2.8d). At scales between 10 and 40 km, the KE spectrum of velocities computed from SSH is essentially white noise, because geostrophy does not hold for a significant fraction of the flows at these scales. Daily-averaging the SSH fields suppresses ageostrophic motions, and the KE spectrum of velocities diagnosed from SSH approximately follows a κ^{-3} power law, consistent with the spectrum of daily-averaged velocities.

The shape of the KE spectra and the partition of the flow into rotational and divergent components show little depth dependence in the upper 1000 m. The main difference is that KE spectra at 1000 m (not shown) lack the sharp peak in the spectra of divergent motions at 100 km. The spectra of daily-averaged fields show a small, but statistically insignificant, depth dependence (the spectra steepens at depth, see figure 2.8c).

Similar conclusions are drawn from SSH variance spectra (Figure 2.9). The variance spectrum of hourly SSH fields is roughly consistent with predictions of isotropic interior QG turbulence at scales between 40 and 150 km. The SSH variance spectrum dramatically flattens at scales smaller than 40 km. Removing high-frequency flows by

daily-averaging the SSH fields suppresses high-wavenumber variance (Figure 2.9). The spectra of daily-averaged SSH approximately follow a κ^{-5} power law between 10 and 150 km.

To assess the fraction of SSH variance due to coherent tides, we remove tidal frequencies using t-tide [Pawlowicz, Beardsley, and Lentz 2002]. In Drake Passage, tidal motions dominate the SSH variability, accounting for 50-80% of the SSH variance. But the majority of this variability is due to barotropic tides, which project onto scales larger than 200 km. There is insignificant tidal variability at scales smaller than 40 km (Figure 2.9). Similar results are obtained by de-tiding the velocity fields (not shown). Hence most of the high-wavenumber ageostrophic variability is likely due to a continuum of inertia-gravity waves, and other ageostrophic flows.

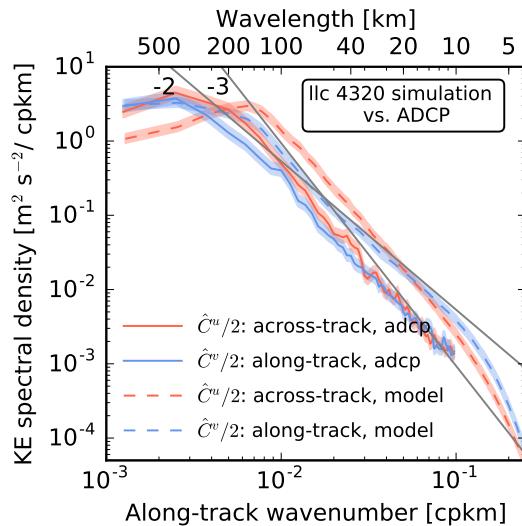


Figure 2.6: Comparison between ADCP and model spectra for the 58-98 m layer. For reference, k^{-2} and k^{-3} curves are plotted (gray lines).

2.7 Discussion

In strong baroclinic currents, the stretching term associated with the mean flow vertical shear dominates the quasigeostrophic potential vorticity (QGPV) gradient [e.g. Tulloch et al. 2011]. Background available potential energy is converted into eddy KE through baroclinic instability locally in wavenumber/frequency space (the fastest

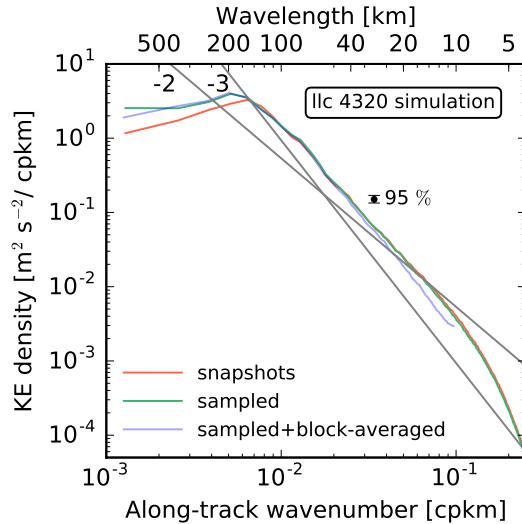


Figure 2.7: Model KE spectra at the surface. The figure depicts spectra for hourly (red), and the “LMG sampled” and block-averaged fields (yellow and green line). For reference, k^{-2} and k^{-3} curves are plotted (gray lines).

growing mode dominates the conversion). In QG turbulence, the KE spectrum is then filled through triad interactions that result in an inverse energy cascade and a direct enstrophy cascade [e.g. Salmon 1998].

Based on a local, linear QG stability analysis of climatological fields, Tulloch et al. [2011] classified both the ACC and the Gulf Stream in the same regime. This regime is characterized by a deep reversal in the QGPV gradient that drives interior QG baroclinic instabilities. In Drake Passage, we estimate the mean deformation radius from cDrake in situ observations [Chereskin, Donohue, and Watts 2012] to be ~ 16 km, and the wavelength of the most unstable modes to be about 188 km (see appendix 2.E). This is roughly consistent with estimates of Tulloch et al. [2011] based on very smooth climatologies, which suggest that the most unstable mode is about twice the deformation scale: $L_{inst} = 2 \times 2\pi \times 16 \approx 200$ km. This forcing scale (or injection scale) is roughly the lowest wavenumber in the range of scales analyzed in our study.

Indeed, the emerging picture of the ACC in Drake Passage has similarities to the one drawn from the ADCP observations in the Gulf Stream [Wang et al. 2010, CF13]. As in the Gulf Stream, the spectral slopes, $2 < n < 3$, are reminiscent of predictions of isotropic QG turbulence in the enstrophy inertial range [Wang et al. 2010, CF13].

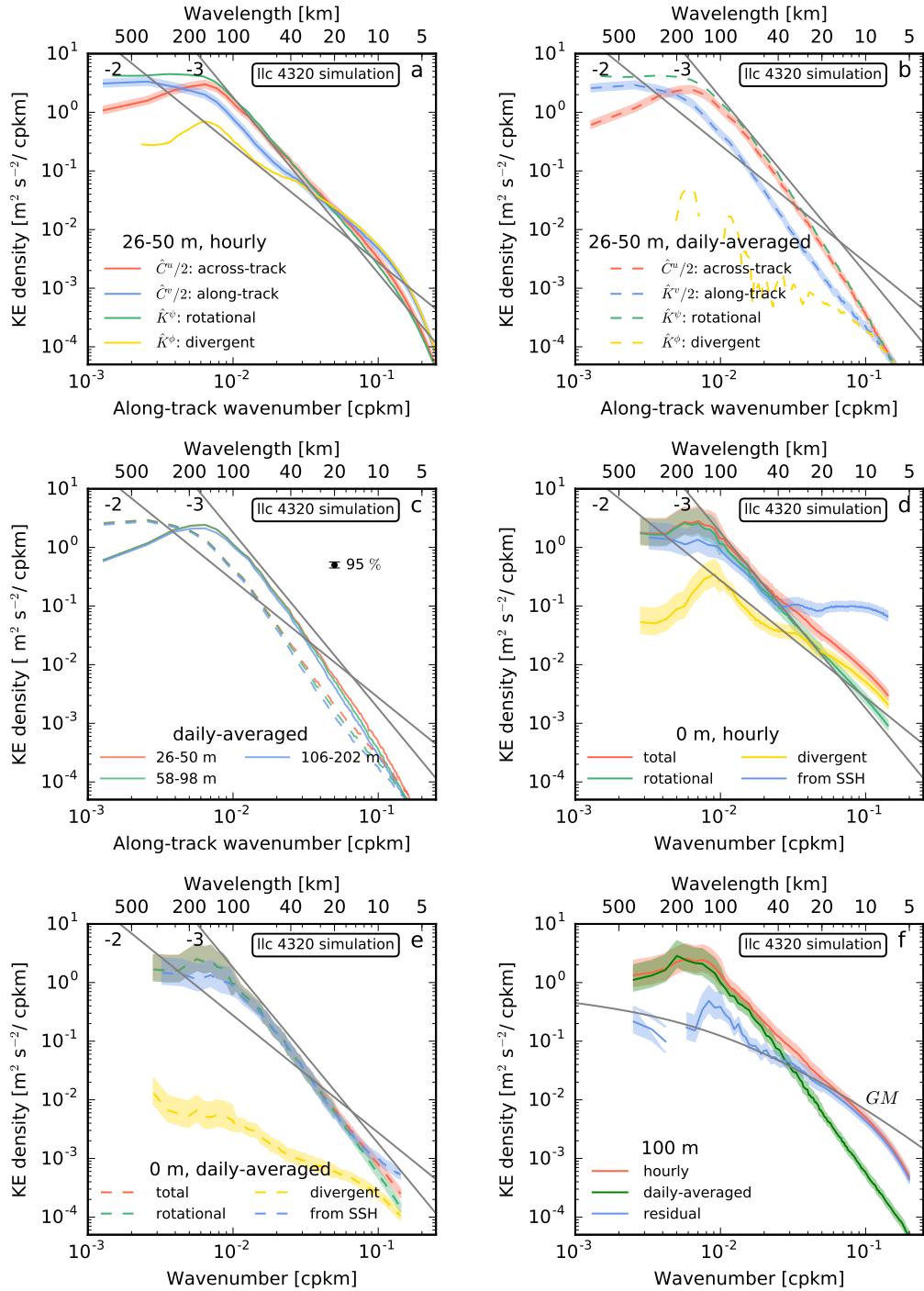


Figure 2.8: Model along-track wavenumber KE spectra for the 25-60 m layer: (a) hourly, (b) daily-averaged; (c) comparison of KE spectra of daily-averaged fields for three layers (solid lines represent across-track spectra, dashed lines along-track spectra); and model isotropic KE spectra at the surface: (d) hourly and (e) daily-averaged; (f) comparison of residual spectrum with the Garrett-Munk spectrum at 100 m. Shaded represent 95 % confidence limits. For reference, and k^{-2} and k^{-3} curves are plotted; in (f) the Garrett-Munk spectrum is plotted with 1.4 times the energy level in Munk [1981].

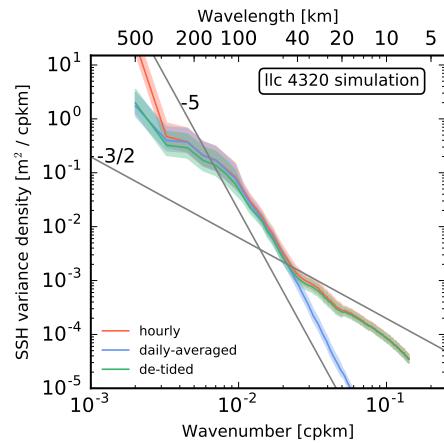


Figure 2.9: Isotropic wavenumber SSH variance spectrum. Shades represent 95% confidence limits. For reference, $k^{-3/2}$ and k^{-5} curves are plotted (gray lines).

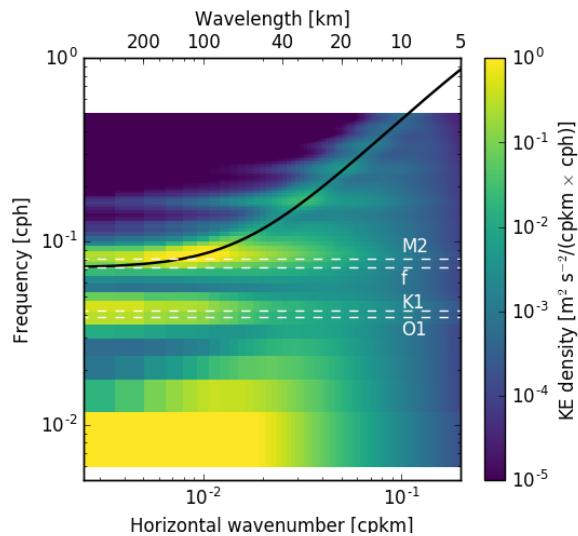


Figure 2.10: Horizontal wavenumber frequency spectrum of KE at 100 m. The solid line is the dispersion relationship for inertia-gravity waves integrated over vertical wavenumber, and the dashed lines indicate tidal and inertial frequencies. The inertial frequency is based on the mean latitude of the domain (about 59°S).

The ratio of across-track to along-track KE energies (R) significantly departs from the expectations of isotropic interior QG turbulence [$R = n$, Charney 1971, CF13] at scales as large as ~ 150 km. The inconsistency between theoretical prediction and the ADCP KE spectra is dramatic in the 10-40 km subrange where we find $R \approx 1$.

The decomposition of the KE spectra into horizontally rotational and divergent components in conjunction with analyses of stochastically-generated velocity fields and data from the llc4320 simulation indicate that horizontally divergent flows are responsible for the difference between R and n . This interpretation is consistent with the analyses of the ADCP data in physical space: the eccentricity of variance ellipses is typically small [Lenn et al. 2007], and thus it is unlikely that the fact that $R < 3$ stems from anisotropies in the flow. However, Lenn et al. [2007] interpret the fact that the ratio of along-passage to across-passage (their coordinate system that is roughly across-track/along-track) velocity variance spectrum is larger than 1 as evidence for anisotropy. While anisotropies may also lead to $R \neq 1$, differences in one-dimensional KE spectra are expected in isotropic QG turbulence [Charney 1971, CF13]. If large-scale eddy flows were markedly anisotropic in Drake Passage, this characteristic would remain in the daily-averaged fields. Instead, model spectra of daily-average velocity are consistent with predictions of isotropic interior QG turbulence (e.g. CF13).

The SSH variance spectrum from the AltiKa altimeter is consistent with ADCP and model spectra at scales larger than about 70 km. At scales between 70 and 200 km, the AltiKa spectrum in Drake Passage is consistent with Topex-Poseidon/Jason spectra in other high eddy kinetic energy regions [Le Traon et al. 2008; Xu and Fu 2012]. We interpret this steep spectral slope as consistent with predictions of isotropic interior QG turbulence. The change in slope in the AltiKa spectrum at scales between 30 and 70 km is consistent with the shallowing of the model SSH variance spectrum due to ageostrophic flows (see figures 2.4 and 2.9). However, the AltiKa spectral flattening should be interpreted with care at these small scales. We were unable to distinguish between noise and ageostrophic flows based only on along-track snap shots of SSH.

In Drake Passage, ageostrophic motions appear to be dominated by inertia-gravity waves. The wave-vortex decomposition of Bühler, Callies, and Ferrari [2014] using the Garrett-Munk spectrum of inertial-gravity waves yields consistent results.

Of course, the validity of this decomposition depends strongly on the assumptions that the Garrett-Munk spectrum accurately depicts the wave fields in the Southern Ocean. Previous studies suggest a rough universality of the Garrett-Munk spectrum [Munk 1981] and the residual numerical model isotropic KE spectrum (hourly KE spectrum minus daily-averaged spectrum) is also roughly consistent with the shape of the Garrett-Munk spectrum (Figure 2.9). Furthermore, a significant fraction of the high wavenumber KE is spread about the dispersion relationship of linear inertia-gravity waves (Figure 2.10). Thus, while all these assumptions are not strictly true, it is unlikely that the Bühler, Callies, and Ferrari [2014] decomposition is misleading, particularly in the low and high wavenumber ends of the spectrum where geostrophic flows and inertia-gravity waves dominate, respectively.

The Garrett-Munk spectrum does not account for tidal peaks. We can estimate the horizontal scale of the nth-mode internal tide using the hydrostatic linear inertia-gravity wave dispersion relationship:

$$L_n \sim 2\pi \times R_n / \sqrt{\omega^2/f_0^2 - 1}, \quad (2.2)$$

where R_n is the nth deformation radius; ω and f_0 are the tidal and the inertial frequencies. For the 1st mode M2 tide $\omega \approx 1.41 \times 10^{-4}$ rad s⁻¹, and using the mean inertial frequency $f_0 \approx 1.25 \times 10^{-4}$ rad s⁻¹, we obtain ~ 146 km and ~ 195 km for model and cDrake stratification, respectively. In the model, this is the peak in the KE spectra of horizontally-divergent motions (the solid yellow line in Figure 2.8a) and peak in the residual spectrum (the solid blue line in Figure 2.8f).

The low-mode internal tides are significantly anisotropic, and the Bühler, Callies, and Ferrari [2014] decomposition may not be very accurate at that peak. In the model, the low mode internal tides are generated on the continental slope off the southern tip of South America. An animation of vertical velocity shows that the low-mode internal tides are significantly scattered and distorted as they propagate into the passage, generating small-scale inertia-gravity waves that are more isotropic. Indeed, tidal flows account for an insignificant fraction of the ageostrophic variability at scales smaller than 40 km.

That ageostrophic motions may contribute significantly to the upper ocean variability should not be a surprise [e.g. Ferrari and Wunsch 2010; Wunsch 2013]. Analysis of a global model with embedded tides suggests that high-frequency motions signifi-

cantly flatten the near-surface spectra [Richman et al. 2012]. By comparing spectra of high-pass and low-pass fields in different regions, Richman et al. [2012] argue that internal tides dominate the ageostrophic signal in SSH; the effect of ageostrophic motions are dramatic in well-known internal tide hot spots [e.g. in the vicinity of the Hawaiian archipelago, Richman et al. 2012]. The present study indicates that ageostrophic flows also contribute significantly to the upper ocean variability in high eddy KE regions. Moreover, the ageostrophic signal is largely nontidal at scales smaller than 40 km.

Interior geostrophic turbulence and inertia-gravity waves coexist at scales between 10 and 200 km, with the former dominating at scales larger than 100 km and the latter dominating at scales smaller than 40 km. In both ADCP and model KE spectra there is no plateau suggesting separation of QG turbulence from ageostrophic flows. Indeed, there is an approximate equipartition of the KE between vortex and wave components between 40 and 100 km. A gap could appear in the vertical wavenumber spectra of other quantities such as shear and strain variance [Pinkel 2014]. The extent to which inertia-gravity waves and geostrophic flows interact to generate the observed spectra deserves further investigation.

There is growing observational evidence for the importance of other mixed-layer intensified motions that develop in response to mesoscale straining; these mixed-layer submesoscale motions are ubiquitous features of the upper ocean [e.g. Shcherbina et al. 2013]. These flows scale roughly with mixed-layer deformation radius [Thomas, Tandon, and Mahadevan 2008]. In Drake Passage, we estimate $L_{ml} = N_{ml}H_{ml}/|f_0| = 100 \times 2.8 \times 10^{-3}/1.3 \times 10^{-4} \approx 2.3$ km, where H_{ml} is the mean mixed-layer depth and N_{ml} is the mean stratification frequency in the mixed-layer from Argo climatology [dataset produced by Roemmich and Gilson 2009]. Given this scale, it is likely that mixed-layer submesoscales are only marginally, if at all, resolved by the ADCP data and the llc4320 simulation. Moreover, the lack of statistically significant depth dependence in the shape of both model and ADCP spectra and in the partition into rotational and divergent components suggest that mixed-layer-confined flows are sub-dominant in the range of scales considered here.

As discussed in Section 2.3, the ADCP data do not fully resolve the horizontal spectrum of inertia-gravity waves. While the exercise of sampling the model results to

mimic the ADCP data suggests insignificant aliasing, block-averaging the data slightly reduces the variance at high wavenumbers, and inertia-gravity waves may be only marginally resolved by the ADCP data (Section 2.6). While block-averaging the ADCP data every 5 km may slightly steepen the KE spectra, the spectra would be unlikely to change from $k^{-5/3}$ to k^{-3} . Indeed, the model spectra are inconsistent with SQG turbulence, which predicts flat KE spectra at the surface ($k^{-5/3}$) and steeper spectra at deeper levels (CF13; see also appendix 2.F).

Finally, the upper ocean is also host to many other flows. For instance stratified turbulence may be important in the pycnocline, whereas Langmuir cells may take place in the upper part of the mixed layer. These “sub-submesoscale” flows (< 1 km) are not resolved by our dataset. The extent to which these flows change the high wavenumber end of the submesoscale range remains to be investigated.

2.8 Conclusions

We remind the reader of the limitations of this study. Ocean turbulence theories, the Bühler, Callies, and Ferrari [2014] decomposition, and the Garrett-Munk spectrum assume stationary, homogeneous, and horizontally isotropic statistics; we also make the first two assumptions when averaging multiple realizations to produce the spectral estimates. In applying the second step of the Bühler, Callies, and Ferrari [2014] decomposition, we assumed that the Garrett-Munk spectrum provides an accurate representation of the inertia-gravity wave spectrum in the Southern Ocean. None of these assumptions are strictly true in the real ocean. Furthermore, the formal, statistical error bars shown with the spectra do not account for unresolved variance in the ADCP or altimeter data, nor do they allow for missing physics in the model. Nevertheless, this study provides the first robust description of the upper ocean mesoscale to submesoscale horizontal wavenumber spectra in the Southern Ocean and an attempt to explore its partition into geostrophic flows and inertia-gravity waves. Our main conclusions are:

- At scales between 10 and 200 km, the upper ocean (0-200 m) horizontal wavenumber spectra in Drake Passage are consistent with isotropic interior QG turbulence masked by unbalanced motions.

- AltiKa altimeter data do not fully resolve all scales from 10 to 200 km, but are consistent with ADCP spectra for scales larger than about 70 km.
- In Drake Passage, ageostrophic motions account for about half of the near-surface kinetic energy at scales between 10 and 40 km, and are likely dominated by inertia-gravity waves.
- Model results indicate that ageostrophic motions significantly imprint on the surface, accounting for about half of the sea surface height variance at scales between 10 and 40 km.

Isotropic interior QG turbulence theory appears to be a relevant framework near strong baroclinic currents [Wang et al. 2010, CF13, this study], although it is only part of the governing dynamics at mesoscales to submesoscales. A robust assessment of the eddy variability in terms of turbulence theories, however, requires evaluation of the transfers of energy and enstrophy across different scales [e.g. Richman et al. 2012]. The present observations prohibit estimation of such diagnostics.

The present results have implications for high-resolution altimeters, e.g., the upcoming Surface Water and Ocean Topography (SWOT) mission. Unbalanced flows may upset the estimation of $\mathcal{O}(10)$ km-resolution surface geostrophic velocities with low noise-to-signal ratios. That is, the high-wavenumber sea surface height variability may represent a different, ageostrophic, physical regime. This problem deserves further investigation.

Appendix 2.A: ADCP error budget

The accuracy in the ADCP relative velocity is estimated as [RDInstruments 1989]:

$$\sigma_{rel} = \frac{1.6 \times 10^5}{F \times D \times \sqrt{N}}, \quad (2.3)$$

where F is the frequency of the sonar, D is the size of the vertical bin, and N is the number of pings averaged to form the relative velocity estimate. For the LMG ADCP, $F = 153.6$ kHz and $D = 8$ m. The ADCP sampling rate is 1 Hz; thus $N = 300$ for 5-min averages. Hence $\sigma_{rel} \sim 7.5 \times 10^{-3}$ m s $^{-1}$.

The GPS system has an accuracy of ~ 5 m. Averaging over $\Delta t = 5$ min yields an estimate for the error stemming from position accuracy of $\sigma_{pos} = \sqrt{2} \times 5/\Delta t \approx 2.4 \times 10^{-2}$ m s $^{-1}$. Hence, the random error in the 5-min absolute horizontal velocity estimates is

$$\sigma_{abs} = \sqrt{\sigma_{rel}^2 + \sigma_{pos}^2} \approx 2.5 \times 10^{-2} \text{ m s}^{-1}. \quad (2.4)$$

Block-averaging the 5-min data every 5 km (~ 17 min for a mean ship velocity $V_{ship} = 10$ kt ≈ 5 m s $^{-1}$) decreases the error as $1/\sqrt{17/5} \approx 0.54$. Thus, the instrumental error in the 5-km absolute velocity is $\sim 1.35 \times 10^{-2}$ m s $^{-1} \sim \mathcal{O}(1)$ cm s $^{-1}$.

Besides random error, there is also a potential bias stemming from errors in the gyroscope. This error is corrected when processing the raw data in CODAS using the GPS heading that appears to be accurate to $< 0.1^\circ$ [see Appendix 3 of Firing, Chereskin, and Mazloff 2011]. An estimate of this bias is then $\sigma_{bias} = V_{ship} \times \sin(0.1^\circ) \approx 7.2 \times 10^{-3}$ m s $^{-1} \sim \mathcal{O}(1)$ cm s $^{-1}$. It is likely that this bias does not significantly affect spectral slopes and the ratio of across-track to along-track KE spectra.

Appendix 2.B: Confidence limits in spectral estimates

To estimate confidence limits (CLs), we assume that the Fourier coefficients are normally-distributed, and consequently that their magnitudes squared are χ^2 -distributed [e.g., Bendat and Piersol 2011]. The average separation between the end of one transect and the start of the next is ~ 13 days. This is about the velocity decorrelation

time scale in Drake Passage as estimated from deep moorings [Bryden 1979]. In the upper ocean, velocity decorrelates faster; the integral time scales estimated from the model are < 8 days. Given that less than 30% of consecutive transects are separated by less than 6 days, and only 10% by less than 4 days, we consider each transect as an independent realization of the spectrum Lenn et al. [see also 2007]. We consider only one independent realization per transect within each layer. Employing these criteria yields $N = 116, 119$, and 55 independent realizations of the spectrum for the 26-50 m, 58-98 m, and 106-202 m layers, respectively; gaps are ubiquitous in deeper bins. For altimetry, we assume each satellite pass to be independent. We consider 32 ascending and 32 descending ground pass segments and make use of the first 20 cycles of the AltiKa mission, and after removing satellite passes with flagged data, this results in $N = 518$ for descending ground tracks and $N = 519$ for ascending ground tracks.

To calculate the number of independent realizations of one-dimensional llc4320 simulation spectral estimates, we first estimate time and length integral scales from autocorrelation functions for daily averaged fields [e.g., Bendat and Piersol 2011]. We obtain, on average, ~ 34 and ~ 45 km for decorrelation length scales of zonal and meridional components, respectively; the mean time scales are ~ 7 and ~ 6 days. Using the largest/longest spatial/time scale yields 200 independent realizations of the one-dimensional spectrum.

As for the isotropic spectra, considering the largest time scale yields 10 independent realizations of the 2D spectra. Given the redness of the spectra, this is likely an underestimate for relatively small scale motions; high wavenumber motions are typically faster than low wavenumber ones. (The exception are tides, which have relatively low horizontal wavenumbers but evolve relatively fast.) We then further assume that motions in the 20-100 and 10-20 km subranges decorrelate on 2.5 and 1 days, respectively. These criteria yield 10, 40, 80 independent realizations of the isotropic spectra for above 100, 20-100, and 10-20 km subranges, respectively.

Appendix 2.C: Details of the Helmholtz and wave-vortex decompositions of one-dimensional KE spectra

To decompose the across-track ($\hat{C}^u/2$) and along-track ($\hat{C}^v/2$) kinetic energy spectra into rotational (\hat{C}^ψ) and divergent (\hat{C}^ϕ) components, we first compute the spectral functions \hat{F}^ψ and \hat{F}^ϕ by solving the coupled system [Bühler, Callies, and Ferrari 2014]

$$k \frac{d\hat{F}^\phi}{dk} - \hat{F}^\psi = -\frac{\hat{C}^v}{2}, \quad \text{and} \quad k \frac{d\hat{F}^\psi}{dk} - \hat{F}^\phi = -\frac{\hat{C}^u}{2}, \quad (2.5)$$

The system (2.5) can be solved analytically, imposing that all spectra and spectral functions vanish at $k = \infty$ [Bühler, Callies, and Ferrari 2014]

$$\hat{F}^\psi(s) = \frac{1}{2} \int_s^\infty [\hat{C}^u(\sigma) \cosh(s-\sigma) + \hat{C}^v(\sigma) \sinh(s-\sigma)] d\sigma, \quad (2.6)$$

and

$$\hat{F}^\phi(s) = \frac{1}{2} \int_s^\infty [\hat{C}^u(\sigma) \sinh(s-\sigma) + \hat{C}^v(\sigma) \cosh(s-\sigma)] d\sigma, \quad (2.7)$$

where we made the following change of variables

$$s \stackrel{\text{def}}{=} \log k, \quad \text{and} \quad \frac{d}{ds} = k \frac{d}{dk}. \quad (2.8)$$

Contrary to Bühler, Callies, and Ferrari [2014], we use u (not v) for the across-track velocity component.

We evaluate the integrals (2.6) and (2.7) numerically using Simpson's rule. Of course, we do not know \hat{C}^u and \hat{C}^v at very small scales, and we stop the integration of (2.6) and (2.7) short of ∞ . Hence, the Bühler, Callies, and Ferrari [2014] decomposition is not very accurate at very high wavenumbers. The rotational (\hat{C}^ψ) and divergent (\hat{C}^ϕ) spectra are computed from the spectral functions \hat{F}^ψ and \hat{F}^ϕ and the KE spectra \hat{C}^u and \hat{C}^v [Bühler, Callies, and Ferrari 2014]

$$\hat{C}^\psi(k) = \left(1 - k \frac{d}{dk}\right) \hat{F}^\psi = \hat{F}^\psi - \hat{F}^\phi + \frac{\hat{C}^u}{2}, \quad (2.9)$$

and

$$\hat{C}^\phi = \left(1 - k \frac{d}{dk}\right) \hat{F}^\phi = \hat{F}^\phi - \hat{F}^\psi + \frac{\hat{C}^v}{2}. \quad (2.10)$$

For the wave-vortex decomposition, we use the Garrett-Munk spectrum [Munk 1981] to obtain the ratio $\hat{F}_W^\psi(k)/\hat{F}_W^\phi(k)$ for the inertia-gravity waves [Bühler, Callies, and Ferrari 2014]. Munk [1981] gives an empirical expression for the total energy spectrum

$$E_{GM}(\omega, j) = E_0 B(\omega) H(j), \quad (2.11)$$

where E_0 is a constant,

$$B(\omega) \stackrel{\text{def}}{=} C \frac{2f}{\pi\omega} (\omega^2 - f_0^2)^{-1/2}, \quad \int_f^{N(z)} B(\omega) d\omega = 1, \quad (2.12)$$

where C is a normalization constant, and

$$H(j) = \frac{(j^2 + j_*^2)^{-1}}{\sum_1^\infty (j^2 + j_*^2)^{-1}}. \quad (2.13)$$

Note that Munk (1981) assumes an exponential stratification profile with e-folding vertical scale h

$$N(z) = N_0 e^{z/h}, \quad (2.14)$$

where N_0 is the buoyancy frequency at the base of the mixed layer, and approximates the WKB vertical wavenumber by

$$m(z) \approx \frac{\pi}{h} \frac{N(z)}{N_0} j, \quad (2.15)$$

Also in (2.12) ω is the frequency, related to the horizontal and vertical wavenumbers by the dispersion relationship

$$\omega^2(k_h^2 + m^2) = k_h^2 N^2 + m^2 f_0^2. \quad (2.16)$$

We can write the GM spectrum as a function of the horizontal wavenumber k_h using the dispersion relationship

$$E_{GM}(k_h, j) = E_{GM}(\omega, j) \frac{d\omega}{dk_h}. \quad (2.17)$$

To calculate the ratio $\hat{F}^\phi / \hat{F}^\psi$ we first notice that plane-wave inertial-gravity waves obey simple a relationship between the total energy spectrum $E(k_h, m)$ and the spectrum of

the horizontal velocity potential ϕ , associated with the horizontally divergent part of the flow. The kinetic energy spectrum of inertia-gravity waves is

$$\hat{K}_W(k_h, m) = \left(1 + \frac{f_0^2}{\omega^2} + \frac{k_h^2}{m^2} \right) k_h^2 \hat{C}_W^\phi(k_h, m), \quad (2.18)$$

where we used $\hat{C}^\psi = \frac{f_0^2}{\omega^2} \hat{C}^\phi$. The potential energy is

$$\hat{P}_W(k_h, m) = \frac{N^2}{\omega^2} \frac{k_h^4}{m^2} \hat{C}_W^\phi(k_h, m). \quad (2.19)$$

Adding (2.18) to (2.19), and using the dispersion relationship (2.16), we obtain

$$\hat{E}_W(k_h, m) = \left(1 + \frac{k_h^2}{m^2} \right) k_h^2 \hat{C}_W^\phi(k_h, m), \quad (2.20)$$

Making the hydrostatic approximation (i.e., ignoring the k_h^2/m^2 above) results in an error of less than 1% at scales larger than 10 km. To obtain $\hat{F}_W^\phi(k)$ we sum on j , and transform the spectra from two-dimensional, horizontally isotropic, to one-dimensional spectrum [e.g., Bühler, Callies, and Ferrari 2014]. We obtain

$$\hat{F}_W^\phi(k) = \frac{1}{\pi} \int_{|k|}^{\infty} \sum_j \frac{E_{GM}(k_h, j)}{k_h^2 + k_h^4/m^2} \sqrt{k_h^2 - k^2} dk_h, \quad (2.21)$$

Similarly,

$$\hat{F}_W^\psi(k) = \frac{1}{\pi} \int_{|k|}^{\infty} \sum_j \frac{f_0^2}{\omega^2} \frac{E_{GM}(k_h, j)}{k_h^2 + k_h^4/m^2} \sqrt{k_h^2 - k^2} dk_h, \quad (2.22)$$

The integrals in (2.21) and (2.22) are evaluated numerically using Simpsons's rule. Given that the divergent component $\hat{C}^\phi(k)$ is assumed to be only due to inertia-gravity waves, $F^\phi(k) = F_W^\phi(k)$, and a full wave-vortex decomposition is achieved [Bühler, Callies, and Ferrari 2014]

$$\frac{\hat{C}_W^\nu(k)}{2} = \frac{f_0^2}{\omega_*^2(k)} \hat{F}^\phi - \hat{F}^\psi + \frac{\hat{C}^\nu}{2}, \quad (2.23)$$

and

$$\frac{\hat{C}_W^u(k)}{2} = \left[1 - k \frac{d}{dk} \left(\frac{f_0^2}{\omega_*^2(k)} \right) \right] \hat{F}^\phi - \frac{f_0^2}{\omega_*^2(k)} \left(\hat{F}^\psi - \frac{\hat{C}^\nu}{2} \right). \quad (2.24)$$

Equations (2.23) and (2.24) follow from (2.5), and $f_0^2/\omega_*^2(k) \stackrel{\text{def}}{=} \hat{F}_W^\psi/\hat{F}_W^\phi$ [Bühler, Callies, and Ferrari 2014]. Note that only $f_0^2/\omega_*^2(k)$ needs to be differentiated. Finally, the vortex (geostrophic flow) components are obtained by subtraction $\hat{C}_V^u = \hat{C}^u - \hat{C}_W^u$ and $\hat{C}_V^\nu = \hat{C}^\nu - \hat{C}_W^\nu$.

Table 2.2: Details of the llc spin-up hierarchy. In this study, we analyze a subset of the llc 4320 simulation is analyzed.

Simulation	Resolution	Time-step [s]	Period	Tides
ECCO2	1/6°	1200	January 2009 - December 2011	No
llc 1080	1/12°	90	Janurary 2010 - July 2012	No
llc 2160	1/24°	45	Janurary 2011 - April 2013	Yes
llc 4320	1/48°	25	Septmber 2011 - December 2011	Yes

Appendix 2.D: Details of the llc 4320 simulation

The llc4320 simulation is based on a global ocean and sea ice configuration of the Massachusetts Institute of Technology general circulation model [MITgcm, Marshall et al. 1997; Hill et al. 2007] and uses a latitude-longitude-polar cap (llc) grid [Forget et al. 2015]. The MITgcm was spun-up in a hierarchy of increasing horizontal resolutions with 90 vertical levels (details in table D2). The simulation analyzed here is the highest resolution, llc 4320, with a time step of 25 s. The prognostic model variables are saved at hourly intervals. Control files and details of the high-resolution llc model set-ups and forcing files are available at http://mitgcm.org/viewvc/MITgcm/MITgcm_contrib/llc_hires.

Surface fluxes are from the 0.14° European Centre for Medium-range Weather Forecasting (ECMWF) atmospheric operational model analysis, starting 2011. The surface fields include 10-m wind velocity, 2-m air temperature and humidity, downwelling long and shortwave radiation, and atmospheric pressure load. These fields are converted to surface fluxes using the bulk formulae of Large and Yeager [2004] and the dynamic/thermodynamic sea ice model of Losch et al. [2010]. The model also includes tidal forcing for the 16 most significant components that are applied as additional atmospheric pressure forcing (A. Chaudhuri 2014, personal communication).

Appendix 2.E: Linear stability analysis

We perform a standard local and linear quasi-geostrophic stability analysis [e.g., Tulloch et al. 2011] to a base state derived from the cDrake experiment [Chereskin, Donohue, and Watts 2012]. The velocity profile is a 4-year time-mean estimated at a location in the Polar Front in Drake Passage from daily objective maps of geostrophic velocity using a two-dimensional array of Current and Pressure-Recording Inverted Echo Sounders from the cDrake experiment [Firing et al. 2014]. The buoyancy profile is a 4-year time-mean at the same location, from daily maps of acoustic travel time. An empirical lookup table of hydrographic properties constructed from nearly 600 CTD and Argo profiles in the cDrake region is used to convert the mean travel time to a corresponding buoyancy profile [Chidichimo et al. 2014].

The structure of the growth rate curves is roughly aligned with the mean velocity profile (Figure 2.12). The most unstable has a wavelength of about 188 km, with a growth rate of about 0.17 days^{-1} . These instabilities are Phillips-like modes [e.g., Tulloch et al. 2011] associated with a deep sign reversal of the background potential vorticity gradient.

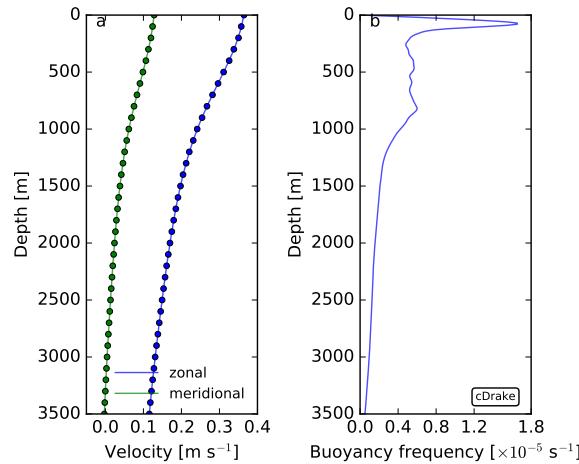


Figure 2.11: Base state used in the linear stability calculations.

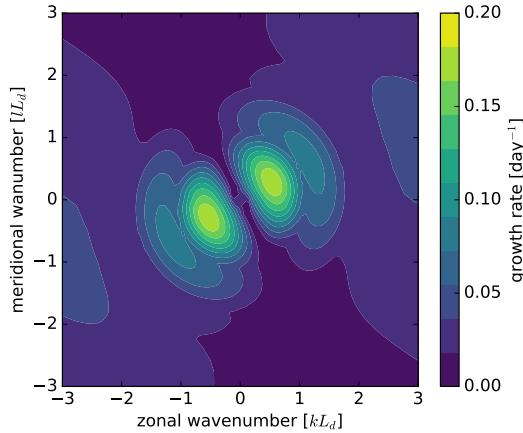


Figure 2.12: Growth rate in horizontal wavenumber space. The wavenumbers are normalized by the deformation scale ($2\pi \times 16$ km).

Appendix 2.F: Depth-dependence of SQG KE spectra for Drake Passage stratification

The vertical structure of surface quasigeostrophic (SQG) solutions $\chi(z; \kappa)$ is governed by [e.g. LaCasce 2012]

$$\frac{d}{dz} \left(\frac{f_0^2}{N^2(z)} \frac{d\chi}{dz} \right) - \kappa^2 \chi = 0, \quad (2.25)$$

subject to inhomogeneous Neumann boundary conditions at the surface that represent surface buoyancy variations

$$\frac{d\chi}{dz} = 1 \quad \text{at} \quad z = 0. \quad (2.26)$$

We assume zero buoyancy anomalies at the bottom

$$\frac{d\chi}{dz} = 0 \quad \text{at} \quad z = -H. \quad (2.27)$$

where $\kappa = \sqrt{k^2 + l^2}$ is the magnitude of the horizontal wavenumber, and $N(z)/f_0$ is the local stratification/inertial frequency.

We generate a SQG-like surface KE spectrum and use the SQG vertical structure $F(z; \kappa)$ to investigate its vertical structure. The SQG vertical structure $F(z; \kappa)$ is computed by solving (2.25)-(2.26)-(2.27) using a second-order finite difference scheme.

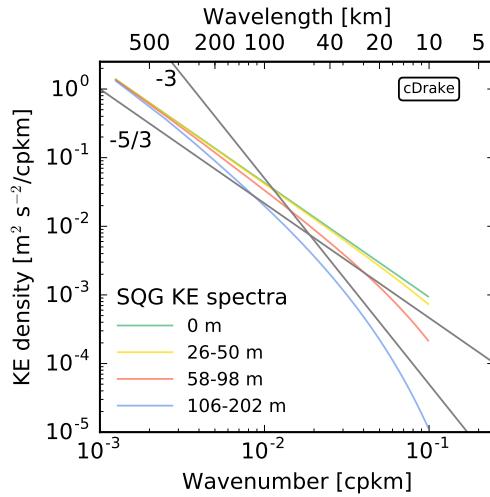


Figure 2.13: Depth dependence of SQG-like KE spectra in the range of scales resolved by the ADCP data. The surface spectra follows a $\kappa^{-5/3}$ power law. The vertical dependence is computed from the SQG vertical structure for each κ given the global average stratification from the cDrake experiment. Colors represent spectra in different layers used in computing the ADCP and llc4320 spectral estimates.

We use the average in situ stratification based on 76 top-bottom CTD profiles from the cDrake experiment [Chereskin, Donohue, and Watts 2012]. As expected, the KE spectrum steepens at depth; the SQG flow at depth is a “low-pass of the surface field” [LaCasce 2012]. For instance, the KE spectra in the 10-50 km subrange approximately follows a power law between κ^{-2} and κ^{-3} in the 106-202 m layer.

In summary, a small (but systematic) steepening in KE spectrum would be present in the upper 200 m, if SQG flows dominated the upper ocean variability in Drake Passage. Both ADCP and model KE spectra lack such systematic depth dependence (see Sections 2.3 and 2.6), and therefore are inconsistent with predictions of SQG turbulence.

Acknowledgments

We had stimulating discussions with William R. Young. Comments by Young, Jörn Callies, Greg Wagner, and two reviewers helped to improve the first draft. Discussions with Callies lead to appendix 2.C. This study was funded by the NASA Ocean Surface Topography Science Team (NNX13AE44G) and SWOT Science Definition Team

(NNX13AE85G), and the NSF Polar Programs (PLR-1341431; PLR-1141922). CBR was partially supported by NSF (OCE 1357047). We are grateful to the captain and crew of the ARSV Laurence M. Gould and to the Antarctic Support Contractor for their excellent technical and logistical support on the cruises. Sharon Escher, Eric Firing and Jules Hummon have been invaluable for their contribution to the ADCP data collection, processing and editing. Caroline Papadopoulos downloaded and archived the model data onto our servers, and Jinbo Wang helped with code for efficiently loading the model data in Python. We are also grateful to C. Hill and G. Forget at MIT; C. Henze and B. Nelson at NASA Advanced Supercomputing (NAS); A. Chaudhuri at Atmospheric and Environmental Research (AER); and the many other MITgcm developers, NAS computer scientists, and SGI engineers that made possible and available the high-resolution model output.

The ADCP and altimeter datasets used in this project are available from the data servers cited throughout the text; the control files for the llc simulations are also available from the MITgcm repository. Plotting scripts, supplementary materials, and data (e.g., ADCP spectra) are openly available at github.com/crocha700/dp_spectra.

Chapter 2, in full, reprints material as it appears in *Journal of Physical Oceanography*, 2016, 46, doi:10.1175/JPO-D-15-0087.1. Rocha, C. B.; Chereskin, T. K.; Gille, S. T.; Menemenlis, D. The dissertation author was the primary investigator and author of this paper.

Chapter 3

Seasonality of submesoscale dynamics in the Kuroshio Extension

3.1 Introduction

Recent interest in upper-ocean dynamics has focused on the strong seasonal cycle of shallow baroclinic instabilities and their role in submesoscale (roughly 1-100 km) turbulence and mesoscale (roughly 100-300 km) modulation [Sasaki et al. 2014; Qiu et al. 2014; Brannigan et al. 2015; Callies et al. 2015; Thompson et al. 2016; Buckingham et al. 2016]. Contemporary studies have also shown that inertia-gravity waves (IGWs) contribute significantly to the near-surface variability at submesoscales [Richman et al. 2012; Bühler, Callies, and Ferrari 2014; Rocha et al. 2016b], but their seasonality has not been investigated.

The partition between geostrophic and ageostrophic flows across the submesoscales and their seasonality have immediate applications for the planning of the Surface Water and Ocean Topography (SWOT) satellite mission. SWOT will deliver sea-surface height (SSH) with about 15 km resolution, thereby providing the first global SSH measurements at submesoscales [Fu and Ubelmann 2014]. Thus, the projection of energetic submesoscale ageostrophic flows, such as inertia-gravity waves (IGWs), onto SSH represents a challenge and an opportunity for the SWOT community. The challenge: how to diagnose submesoscale surface velocity from SSH. The opportunity: stimulation of research on interactions between geostrophic and ageostrophic flows and

their projection onto SSH.

Using the output of two comprehensive global numerical simulations, here we show that IGWs undergo a strong near-surface seasonal cycle in the Kuroshio Extension region. Interestingly, the seasonal cycle of IGWs is out of phase with the seasonal cycle of submesoscale turbulence. Consistent with previous studies, submesoscale turbulence is strongest in winter [Sasaki et al. 2014; Qiu et al. 2014]. IGWs with horizontal scales between 10-100 km (hereafter submesoscale IGWs), however, peak in late summer/early fall, when the upper ocean is strongly stratified. Thus there exists a strong seasonal modulation of upper-ocean submesoscale dynamics: submesoscale turbulence dominates the upper-ocean dynamics in late winter/early spring, whereas submesoscale IGWs prevail in late summer/early fall.

3.2 The LLC numerical simulations

We use results of two latitude-longitude polar cap (LLC) comprehensive numerical simulations. The outputs analyzed here, LLC2160 (nominal resolution $1/24^\circ \approx 4.7$ km; effective resolution ≈ 20 km) and LLC4320 (nominal resolution $1/48^\circ \approx 2.3$ km; effective resolution ≈ 8 km), are forward solutions of the Massachusetts Institute of Technology general circulation model [MITgcm; Marshall et al. 1997] on a LLC grid [Forget et al. 2015] with 90 vertical levels. The coarser-resolution LLC simulation was spun up from an Estimating the Circulation and Climate of the Ocean, Phase II [ECCO2; Menemenlis et al. 2008] adjoint-method state estimate, constrained to millions of observations from 2009 through 2011.

The LLC2160 output spans two years from March 2011 to April 2013; the LLC4320 was spun up from the LLC2160 simulation, spanning one year from September 2011 to October 2012. Both LLC2160 and LLC4320 simulations were forced by tides hourly and by 6-hourly surface atmospheric fields. The LLC4320 simulation is an extension of the 3-month long output used by Rocha et al. [2016b] — for details, see supporting information.

Tidal forcing by the 16 most-significant components represents a key feature of the LLC2160 and LLC4320 simulations. Barotropic tides flowing over topography

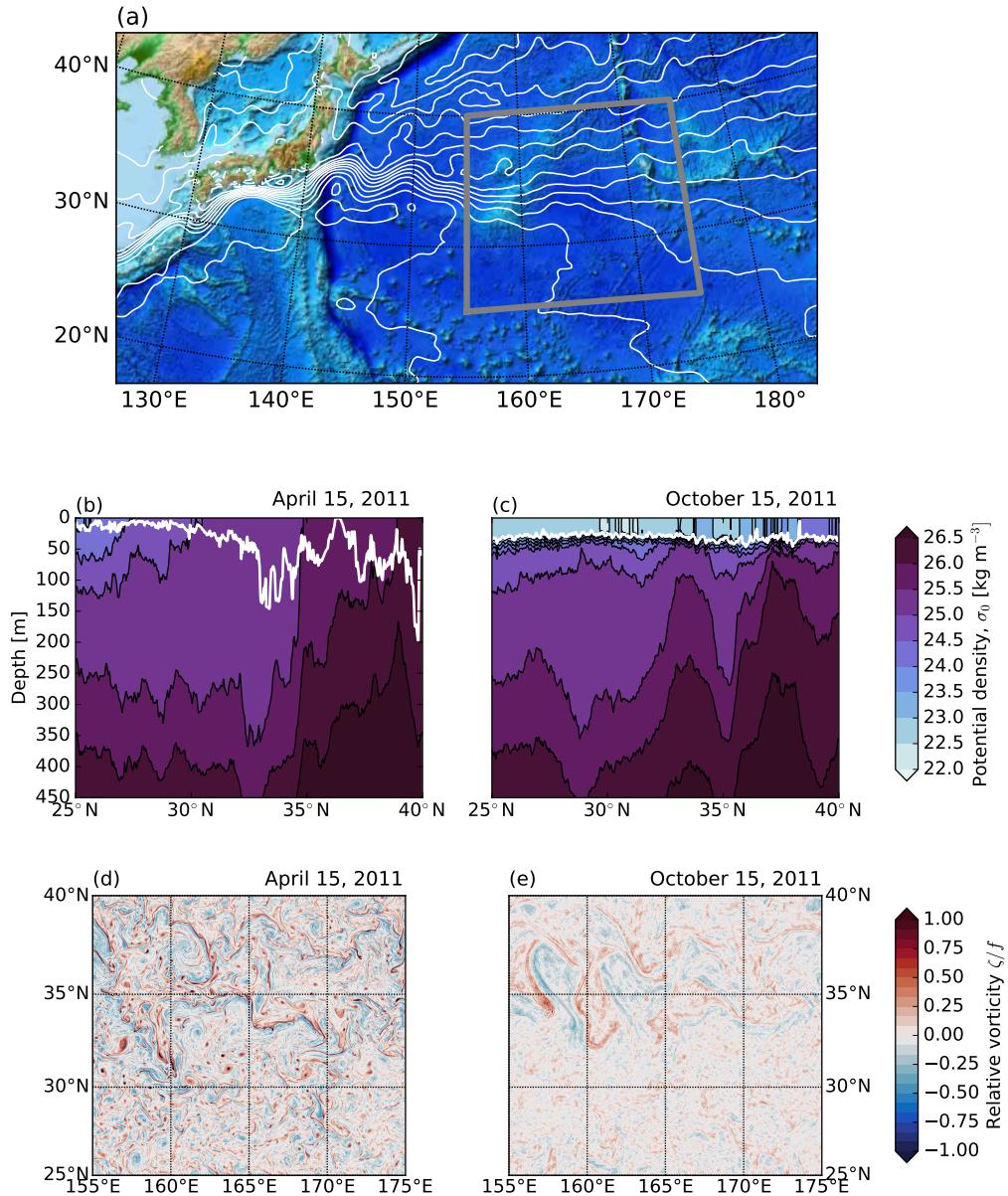


Figure 3.1: (a) The study region with the subregion where the LLC outputs are analyzed. Colors represent the topography and white lines are contours of absolute dynamic topography every 0.1 m from AVISO. LLC 4320 (1/48°) snapshots of transects of potential density at 165°E (b and c) and surface vorticity (d and e). In (b) and (c) the white line represents the mixed layer depth. The snapshots were taken at 12:00 UTC.

convert significant tidal energy into the internal wave field. Those internal waves project onto mesoscales to submesoscales [e.g., Rocha et al. 2016b]. Hence tidal forcing

fundamentally distinguishes our analysis from other modeling studies of seasonality [Sasaki et al. 2014; Qiu et al. 2014].

To study seasonal variations in the upper-ocean dynamics, we focus on the northwest Pacific, in the vicinity of the Kuroshio Extension, where previous studies have shown strong mesoscale and submesoscale seasonality [Sasaki et al. 2014; Qiu et al. 2014]. We analyze a sub-domain of the LLC4320 and LLC2160 simulations of about 2000 km² spanning 155-175°E; 25-40°N (Figure 3.1a). The stratification in this mesoscale-rich subtropical region undergoes a vigorous seasonal cycle: wintertime-enhanced small-scale turbulence de-stratifies the upper ocean, yielding mixed layers deeper than 200 m to the north (Figure 3.1b). Summertime re-stratification yields mixed layers shallower than 40 m (Figure 3.1c).

Fundamentally, the upper-ocean density structure is well-captured by both LLC simulations: a comparison with Argo climatology [updated from Roemmich and Gilson 2009] shows that both simulations skillfully represent the Kuroshio Extension stratification and its seasonal variability (supporting information). Moreover, model smoothed fields have eddy kinetic energy (EKE) levels (0.041 m² s⁻²) similar to Jason II across-track geostrophic EKE (0.036 m² s⁻²) and gridded geostrophic EKE (0.032 m² s⁻²).

3.3 Bulk statistics of the surface lateral velocity gradient tensor

To study the seasonality in the surface velocity, we calculate the lateral velocity gradient tensor

$$\begin{bmatrix} u_x & u_y \\ v_x & v_y \end{bmatrix} \quad (3.1)$$

using a centered second-order finite difference scheme — we obtain qualitatively similar results using spectral differentiation of velocity fields periodized with reflections; statistics of fields obtained via spectral differentiation are 10-15% larger. We then diagnose the vertical vorticity

$$\zeta \equiv v_x - u_y, \quad (3.2)$$

lateral rate of strain

$$\alpha \equiv [(u_x - v_y)^2 + (u_y + v_x)^2]^{1/2}, \quad (3.3)$$

and horizontal divergence

$$\delta \equiv u_x + v_y. \quad (3.4)$$

Because vortices, fronts, and IGWs are ubiquitous at submesoscales, second-order statistics of vorticity, strain, and divergence highlight these fine lateral scales [e.g., Shcherbina et al. 2013].

The model solutions depict seasonality in vorticity (Figures 3.1d-e): large values of fine-grained vertical vorticity are observed in early spring with maximum values as large as $4f$, where f is the local planetary vorticity, and root-mean-square (RMS) of about $0.4f$. In early fall, the situation is the opposite: the vertical vorticity is relatively coarse-grained; its local maximum and RMS are both smaller than $0.5f$. Indeed, the vorticity and rate of strain are strongest in wintertime (Figure 3.2): in both simulations, the RMS vorticity and strain rate are about twice as large in late winter/early spring than in late summer/early fall. Because the wintertime vorticity and strain rate are dominated by the smallest scales in the flow (the KE spectrum is shallower than a -3 power law in winter), increasing the resolution from $1/24^\circ$ to $1/48^\circ$ increases the wintertime RMS vorticity and strain by about 40%.

Subinertial flows account for the bulk of vertical vorticity and strain rate ($T_{32.5} = 2\pi/f_{32.5} \approx 22.3$ h, where $f_{32.5}$ is the inertial frequency at the mid-latitude 32.5°N): daily-averaging the velocity fields suppresses super-inertial motions and reduces the RMS vorticity by 40% and the RMS strain by 10%; the seasonal cycle remains strong (see dashed lines in figures 3.2a-b). Indeed, submesoscale flows account for most of this seasonal cycle: smoothing the velocity fields with a Hanning filter with cut off scale of 100 km dramatically reduces the RMS vorticity and strain rate. The reduction in variance is about 80% in winter, yielding RMS vorticity and strain rate roughly consistent with the diagnostics from AVISO gridded geostrophic velocities (compare red lines to black lines in figures 3.2a-b). The picture that emerges is consistent with recent studies: shallow baroclinic instabilities energize the submesoscales in winter, drawing from the available potential energy stored in large lateral buoyancy gradients in deep mixed layers [Sasaki et al. 2014; Callies et al. 2015; Callies et al. 2016].

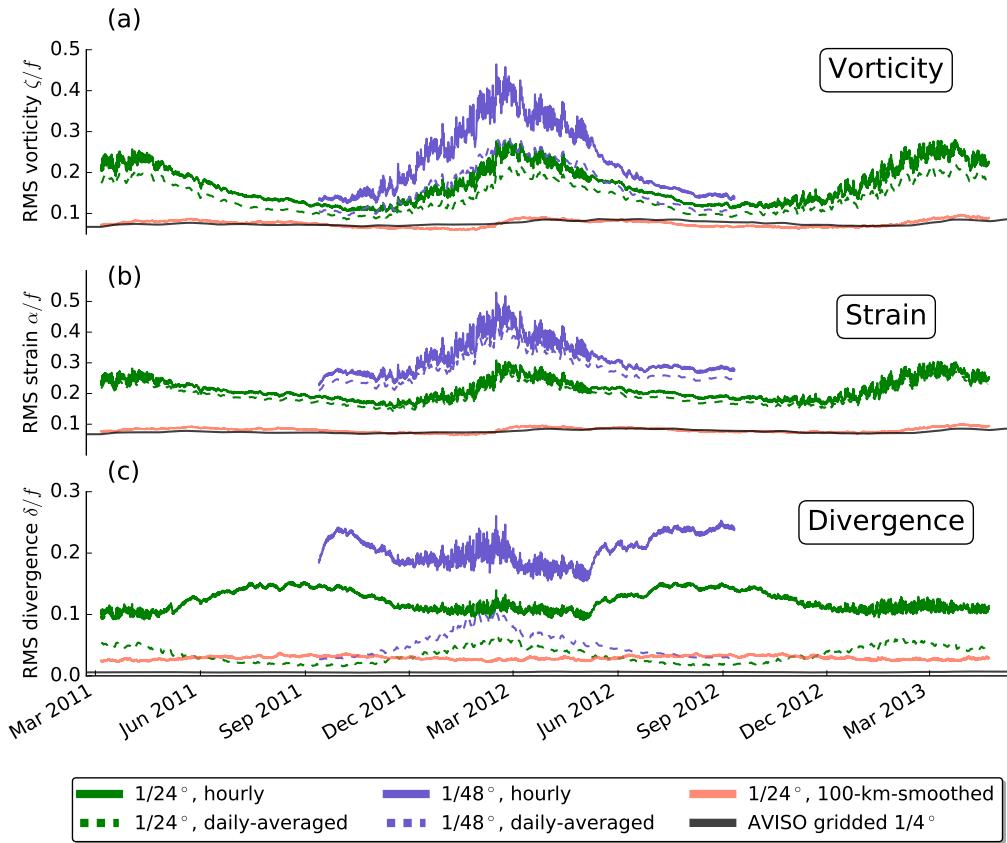


Figure 3.2: Time series of the root-mean-square (RMS) of surface (a) vorticity, (b) rate of strain, and (c) horizontal divergence in the LLC outputs and gridded AVISO data. The convergence of meridians account for the small RMS divergence in AVISO line in (c).

The seasonal cycle of the horizontal divergence, however, showcases the complexity of the upper-ocean annual variability. If submesoscale turbulence dominated the near-surface variability all year, then the seasonal cycle of horizontal divergence, vertical vorticity, and lateral strain rate would be in phase [e.g., Sasaki et al. 2014]. While there is a clear wintertime peak in divergence of daily-averaged velocity (see dashed lines in figure 3.2c; RMS divergence $\sim 0.1f$ in the $1/48^\circ$ simulation), the hourly fields show a stronger enhancement of lateral divergence in late summer/early fall (RMS divergence $\sim 0.22f$ in the $1/48^\circ$ simulation). Because the $1/48^\circ$ simulation better resolves smaller-scale submesoscale flows, a secondary RMS divergence peak in winter is nearly as strong as in summer. Submesoscale fronts and eddies evolve relatively fast, and there is no clear temporal and spatial scale separation between those motions

and IGWs [McWilliams 2016]: daily-averaging the velocity fields efficiently suppresses the summertime horizontally divergent flows, but also reduces the wintertime lateral divergence by about 50%. Figure 3.2c also shows that most of the lateral divergence is associated with submesoscale flows: smoothing the velocity fields with a 100-km-cutoff suppresses more than 80% of the RMS divergence.

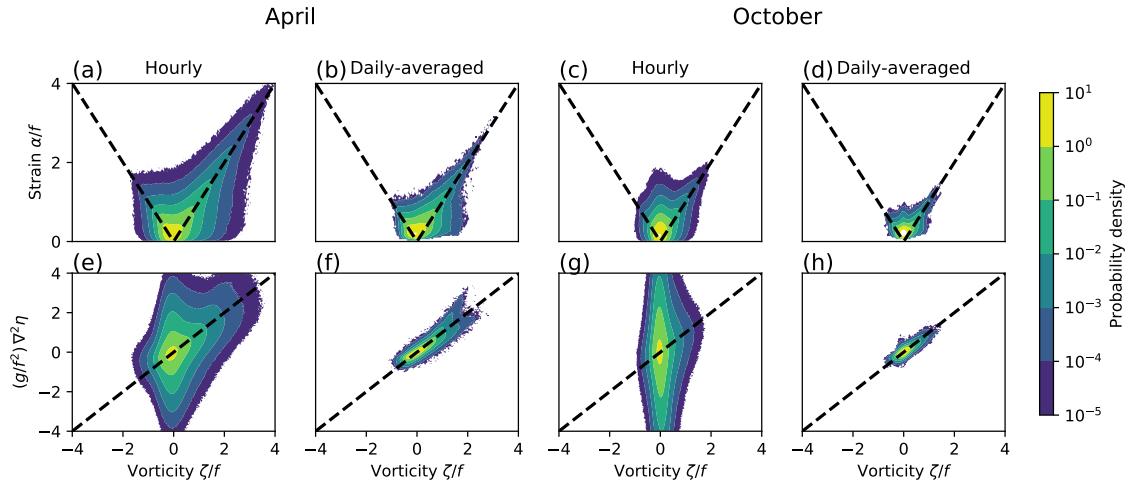


Figure 3.3: Seasonal variation of joint probability distributions of surface lateral velocity gradient tensor: vorticity vs. strain rate (a through d), and vorticity vs. Laplacian of sea-surface height (e through h) in April (a, b, e, f) and October (c, d, g, h). Dashed lines in (a) through (d) represent one dimensional shear flow $\alpha = \pm\zeta$, characteristic of fronts. Dashed lines in (e) through (h) represent geostrophic flow $\zeta = \frac{g}{f} \nabla_h^2 \eta$.

3.4 Joint probability density functions

The results of figure 3.2a-c show that submesoscale surface variability stems from different dynamics in summer than in winter. To characterize these differences, we calculate joint probability distributions (joint-PDF) of vorticity-strain and vorticity-Laplacian of sea-surface height (Figure 3.3).

The April vorticity-strain joint-PDF has a shape characteristic of submesoscale turbulence (Figures 3.3a-b). The alignment of vorticity and strain $\alpha \sim \pm\zeta$ with strong positive skewness are fingerprints of submesoscale fronts [Shcherbina et al. 2013; McWilliams 2016]. The shape of the vorticity-strain joint-PDF is similar for hourly and

daily-averaged fields, although the vorticity skewness reduces from 1.4 to 1.13 from hourly to daily-averaged. The April results indicate that wintertime submesoscale surface velocity is strongly dominated by submesoscale turbulence. The daily-averaged fields are largely in geostrophic balance, but the hourly velocity and sea-surface height fields have an important ageostrophic component as depicted by the joint-PDF of vorticity-Laplacian of SSH winter (Figure 3.3e).

The October vorticity-strain joint-PDF shows much weaker skewness (the vorticity skewness is 0.68 and 0.67 for hourly and daily-averaged velocities). The shape of the vorticity-strain joint-PDF appears to be a combination of two half-ellipses centered about $\zeta = 0$, one with a 45° slope (characteristic of submesoscale fronts that persist in summer) and one with a very steep slope. That the submesoscale dynamics in October are mainly ageostrophic is clearly depicted in the shape of joint-PDF of vorticity-Laplacian of sea-surface height for hourly fields (Figure 3.3g), which is an ellipse aligned in the vertical axis. Daily-averaging the model suppresses the ageostrophic, super-inertial flows, and, therefore, the daily-averaged flow is essentially geostrophic as depicted by the 45°-tilted ellipse in the vorticity-Laplacian of sea-surface height joint-PDF (Figure 3.3h).

Time series of PDFs of vorticity and divergence (supporting information) show a strong oscillation between these two regimes. In late winter/early spring the vorticity is strongly positively skewed, whereas the divergence is moderately negatively skewed (convergence/downwelling) as predicted by frontogenesis [e.g., McWilliams 2016]. In late summer/early fall, the divergence is stronger, but PDFs are much less skewed, consistent with a superposition of random linear IGWs [e.g., Garrett and Munk 1972].

3.5 Wavenumber-frequency spectrum

To confirm that submesoscale IGWs account for most of the super-inertial SSH variance and surface KE, we calculate the wavenumber-frequency spectrum of SSH variance. Focusing on the high-frequency content, we compute the wavenumber-frequency spectrum every 10 days and average the results to obtain a spectral estimate. Before calculating the discrete Fourier transform, we remove linear trends and multiply the

data by a three-dimensional Hanning window. We azimuthally-average the spectrum in wavenumber.

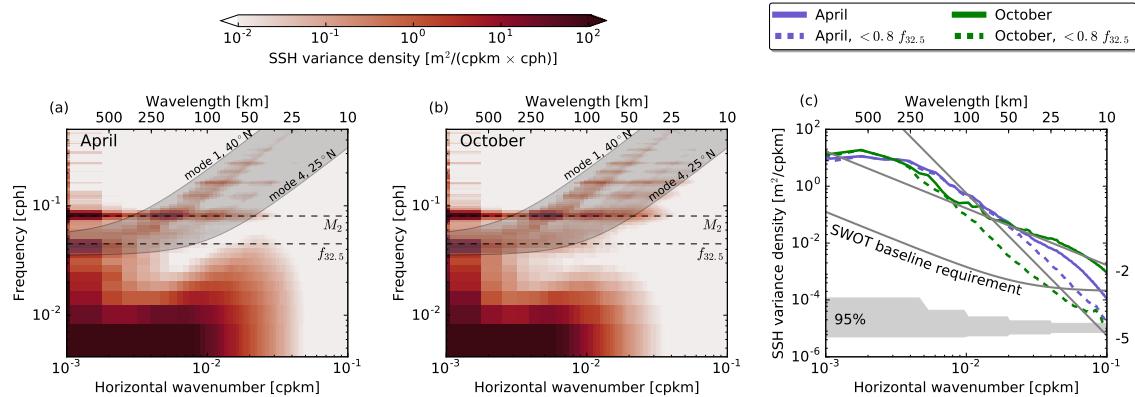


Figure 3.4: LLC4320 wavenumber-frequency spectrum of SSH variance in (a) April and (b) October. (c) Wavenumber spectrum of SSH variance — the integral of (a) and (b) over frequency. In (a) and (b), the light gray shaded region depicts the dispersion relations for inertia-gravity waves from mode 1 through mode 4 across the latitudinal domain; horizontal dashed lines represent the semi-diurnal lunar tidal frequency (M_2) and the inertial frequency at mid-latitude ($f_{32.5}$). For reference, the gray line shows the baseline requirement as specified by the SWOT science team [Rodriguez 2014].

Mesoscales flows and the lunar semi-diurnal tide contain most of the SSH variance (Figures 3.4a-b). But there is significant SSH variance at submesoscales, both sub-inertial and super-inertial. Sub-inertial SSH variance at submesoscales is larger in April, whereas super-inertial SSH variance at submesoscales is larger in October. At scales smaller than 100 km (submesoscales), the SSH variance is 66% larger in October than in April. IGWs, particularly modes two through four, account for most of the SSH variance increase in summer — higher vertical modes are more sensitive to changes in near-surface stratification (supporting information). In October, super-inertial IGWs account for a SSH standard deviation of about 4.9 cm, comparable to the April sub-inertial SSH standard deviation of ~5.2 cm (Table 3.1). Similarly, the submesoscale super-inertial surface KE is 63% larger in October than in April (supporting information).

Figure 3.4c shows the integral over frequency of wavenumber-frequency spectra in figures 3.4a-b. Both in April and October, the SSH wavenumber spectrum approximately follows a -2 power-law at submesoscales (solid lines in figure 3.4c) — there is

Table 3.1: The square root of the mean-square SSH at submesoscales (10-100 km) estimated from the wavenumber-frequency spectrum of SSH variance in figure 3.4. The SWOT baseline requirement at 15 km is approximately $\sqrt{2}$ cm.

	April	October
sub-inertial ($< 0.8f$)	5.2 cm	2.0 cm
super-inertial ($> 1.2f$)	2.9 cm	4.9 cm
near- f ($0.8 - 1.2f$)	0.3 cm	0.4 cm
near- M_2 ($0.8 - 1.2M_2$)	1.5 cm	3.1 cm

small difference between the spectra owing to the phase cancellation between the seasonal cycle of submesoscale turbulence and IGWs. The sub-inertial flow has a steeper submesoscale SSH variance wavenumber spectrum (dashed lines in figure 3.4c). The difference between the SSH variance wavenumber spectrum for all frequencies and sub-inertial frequencies is dramatic in October when super-inertial flows account for about 80% of the SSH variance and KE at submesoscales — we obtain similar results by computing SSH variance and KE wavenumber spectra directly from hourly and daily-averaged SSH velocity fields.

3.6 Discussion and conclusion

Our work adds to recent model [e.g., Sasaki et al. 2014] and observational [e.g., Callies et al. 2015; Buckingham et al. 2016] evidence of vigorous seasonality in submesoscale turbulence. In particular — in global simulations with embedded tides — near-surface submesoscale IGWs undergo a strong seasonal cycle that is out of phase with the seasonal cycle of submesoscale turbulence. IGW energy and SSH variance are significantly spread across the submesoscales, likely through mesoscale advection and refraction, and dispersion [e.g., Ponte and Klein 2015; Alford et al. 2016, and references therein] and parametric sub-harmonic instability [PSI, e.g., MacKinnon and Winters 2005], though the LLC simulations are unlikely to fully resolve PSI. In the LLC simulations, super-inertial IGWs account for most of the surface submesoscale KE and SSH variance; near-inertial IGWs project onto larger scales, possibly because

submesoscale NIWs quickly propagate into the interior and weakly project onto SSH.

D'Asaro [1978] showed that the velocity of linear internal waves in the mixed layer strongly depends on the density jump at the mixed layer base, with largest velocities when the jump is strongest. In summer, the shallow mixed layer overlays a strong seasonal pycnocline, and thus the internal waves projection onto the mixed layer may be stronger according to D'Asaro [1978]'s arguments. An alternative explanation is that the shape of the baroclinic modes changes seasonally [see World Ocean Atlas 2013, Levitus et al. 2013, modal shapes in the supporting information].

Internal tide energy flux does not show a seasonal modulation [e.g., Alford 2003]; hence one expects a strong seasonality in the near-surface expression of internal tides and higher frequency IGWs if the near-surface stratification varies strongly. The intermittent windy near-inertial wave (NIW) generation peaks in early winter, but NIW energy has a similar energy level in October and April [Alford et al. 2016]. Furthermore, in our simulations the higher frequency inertia-gravity waves account for most of the seasonality in super-inertial surface KE and SSH (Figure 3.4a-b).

The seasonally varying mixed-layer depth challenges observation and simulation of flows at mixed-layer deformation-radius length scale in observational programs and models with fixed resolution [Buckingham et al. 2016]. But the mixed layer available potential energy is lowest in summer owing to shallow mixed-layer depths and weak lateral buoyancy gradients [Callies et al. 2015]. Moreover, in baroclinic turbulence, the deformation radius has only a secondary, catalytic role — most of the baroclinic conversion occurs at larger scales [Larichev and Held 1995; Callies et al. 2016]. Thus, the model failure to resolve the summertime ML deformation radius is unlikely to account for the vigorous seasonality in submesoscale turbulence.

In the context of SWOT, these results imply that if one is interested only in the geostrophic flow, so that surface velocity can be diagnosed from SSH, then the noise-to-signal ratio will have a strong seasonality owing to the presence of IGWs (the coherent fraction of internal tides may be removed efficiently). An important caveat is that the Kuroshio Extension may be typical of a mesoscale-rich subtropical regions, but it is unlikely to be representative of other regions such as low-eddy-kinetic-energy eastern boundary currents. We plan to report on the geographic variability of submesoscale

seasonality in a future study.

Acknowledgments

William R. Young and three reviewers provided helpful and stimulating comments. Greg Wagner suggested that the near-surface shape of the baroclinic modes changes seasonally. We thank the MITgcm community and our colleagues at the NASA Advanced Supercomputing (NAS) Division for their awesome support. This research was funded by NSF (OCE1357047) and NASA (NNX13AE44G, NNX13AE85G, NNX16AH67G, NNX16AO5OH). The LLC output can be obtained from the ECCO project (http://ecco2.org/llc_hires). The altimeter products were produced by Ssalto/Duacs and distributed by AVISO, with support from CNES (<http://www.aviso.altimetry.fr/duacs/>). Codes and output files are available online at the project repository (<https://github.com/crocha700/UpperOceanSeasonality>).

Chapter 3, in full, reprints material as it appears in Geophysical Research Letter, 2016, 43, doi: 10.1002/2016GL071349. Rocha, C. B.; Gille, S. T.; Chereskin, T. K.; Menemenlis, D. The dissertation author was the primary investigator and author of this paper.

Chapter 4

Stimulated generation: extraction of energy from balanced flow by near-inertial waves

4.1 Introduction

The inverse cascade, acting on balanced ocean macroturbulence, transfers energy towards large spatial scales. But a statistically steady ocean circulation requires energy dissipation at the same rate as it is supplied by wind. Thus equilibration of the ocean macroturbulence requires ageostrophic processes, acting in opposition to the inverse cascade, to produce a transfer of energy towards the centimeter scales at which molecular viscosity is effective. Mechanisms that might result in this down-scale transfer include, but are not limited to, surface and benthic boundary-layer turbulence, lee-wave generation by mesoscale eddies negotiating bottom topography, and the spontaneous generation of internal waves by upper-ocean frontal instabilities; see Nagai et al. [2015] for a recent review.

We focus on a mechanism first identified by Gertz and Straub [2009]: externally-forced near-inertial waves might provide an energy sink for large-scale balanced flow. Since Gertz and Straub [2009], several other studies, summarized in table 4.1, have argued for significant energy transfer from balanced flows to near-inertial waves. A common aspect of the studies in table 4.1 is that the near-inertial waves are first

Table 4.1: Summary of model-based studies of energy extraction from balanced flow by near-inertial waves.

Study	Framework	Energy transfers from a mean flow to <i>existing</i> near-inertial waves are referred to as
Gertz and Straub [2009]	Barotropic 2D double-gyre solutions coupled with forced 3D near-inertial waves.	<i>2D-to-3D energy transfer</i>
Thomas [2012]	Near-inertial waves in a baroclinic geostrophic flow undergoing frontogenesis.	<i>deformation shear production</i>
Taylor and Straub [2016]	Boussinesq channel-flow with both high- and low-frequency forcing.	<i>advective sink</i>
Barkan, Winters, and McWilliams [2016]	Boussinesq channel-flow with both high- and low-frequency forcing.	<i>direct extraction</i>
Shakespeare and Hogg [2017]	Boussinesq channel-flow with low-frequency forcing. Spontaneous generation in the surface layer and stimulated generation in the interior.	<i>interior amplification</i>

introduced by external forcing (e.g., wind) at the inertial frequency and then grow by extracting energy from the balanced flow.

(An exception in table 4.1 is the study of Shakespeare and Hogg [2017], in which near-inertial waves are generated spontaneously—without external forcing—at density fronts near the surface. These waves then radiate vertically downwards into the interior and amplify by extracting energy from deep balanced flow. We have included Shakespeare and Hogg [2017] in table 1 because, as far as deep interior amplification is concerned, the details of the shallow generation process are probably immaterial.)

The studies in table 4.1 have shown that externally generated near-inertial waves can extract energy from a pre-existing balanced flow. Those studies, however, have diverse methodology and diagnostic framework, so there is not a consensus that the observed amplification of near-inertial waves results from a single mechanism. In other words, it is possible that near-inertial wave amplification occurs through a variety of mechanisms and each instance must be analyzed and understood on its own peculiar terms. But in a certain limit, described in section 4.2.2, there is a single underlying mechanism—stimulated generation—that is responsible for energy transfer between waves and balanced flow (Xie & Vanneste 2015, XV hereafter). While it is unclear whether this limit applies to the studies in table 4.1, XV provides a robust yet simple theoretical framework for studying energy transfers between balanced flows and near-inertial waves.

Using a variational formulation of the generalized Lagrangian-mean, XV derived a phase-averaged model of the coupling between weakly non-linear near-inertial waves and quasi-geostrophic (QG) flow. Wagner and Young [2016] derived a similar coupled model via Eulerian multiple-time expansion; these authors include the second harmonic of the primary near-inertial wave and simplify the wave dynamics by assuming moderate QG vertical shears. In both coupled models the near-inertial waves are governed by the equation of Young and Ben-Jelloul [1997] (YBJ hereafter) and the balanced flow satisfies QG dynamics—the waves contribute phase-averaged quadratic terms at the same order of the QG potential vorticity (PV). Salmon [2016] provides a useful perspective on this “NIW-QG” model; without assuming that the waves are near-inertial, and within a single variational framework, Salmon unifies XV’s model with

the wave-mean flow models of Bühler and McIntyre [1998] and Wagner and Young [2015]. Salmon also emphasizes a revealing analogy between vortex-wave interactions and classical electrodynamics.

To distinguish energy extraction by existing waves from *spontaneous generation*, and to complete an electrodynamic analogy, XV refer to the transfer of energy from balanced flow to externally forced near-inertial waves as *stimulated generation*. The more widely studied process of spontaneous generation is the emission of internal waves arising from the slow evolution of balanced flow in the absence of external forcing at wave frequencies [Vanneste 2013]. Spontaneous generation is inefficient at small and moderate Rossby numbers and its global impact on ocean energetics is probably small [Danioux et al. 2012; Nagai et al. 2015]. And spontaneous generation is localized at sharp submesoscale fronts with order-one Rossby number [e.g., Shakespeare and Hogg 2017] while the stimulated variety operates even at small Rossby numbers characteristic of most interior oceanic conditions, provided only that internal waves are introduced by external forcing. Throughout the ocean, internal waves are reliably forced by wind and tides and thus stimulated generation is a leading contender as a mesoscale energy sink.

Although XV and Wagner and Young [2016] use significantly different approaches, their results are consistent with one another. This consistency indicates that the NIW-QG model provides the unique *small-amplitude* evolution equations describing the interaction between near-inertial waves and geostrophic flow. In the small-amplitude limit the flow can be unambiguously separated into weakly non-linear internal waves and quasi-geostrophic eddies, with perturbative coupling between waves and eddies [Salmon 2016]. To the extent that the studies in table 4.1 are also in this weak-interaction limit, their results should—in principle—be described by the NIW-QG model. “In principle” because the distinction between Lagrangian-mean and Eulerian-mean velocities complicates the diagnosis of energy transfers between eddies and waves (see section 4.6 for further discussion) and because the Rossby number is large in some studies [e.g., Barkan, Winters, and McWilliams 2016]. Frontal sharpening and low Richardson number processes, which are described in Thomas [2012], are outside the scope of the NIW-QG model.

XV emphasize that a central feature of the NIW-QG model is that there are two integral energy conservation laws for: (1) near-inertial kinetic energy and (2) the sum of near-inertial potential energy and total balanced energy. The inevitable reduction of near-inertial length scales by advection and refraction is accompanied by an increase in wave potential energy and, because of conservation law (2), a reduction in the energy of the balanced flow. These features, and the necessity of an externally forced wave, are the defining characteristics of stimulated generation.

Here we investigate perhaps the simplest example of stimulated generation obtained from the NIW-QG model by assuming barotropic QG flow and vertically-planar near-inertial waves. Because the balanced flow is barotropic, while the near-inertial wave is three dimensional, this “vertical-plane-wave model” resembles the original scenario of Gertz and Straub [2009]. We show that the convergence of wave kinetic energy into anti-cyclones and geostrophic straining of the waves reduces the wave length scale, amplifies gradients of wave amplitude, and converts balanced kinetic energy into near-inertial potential energy.

4.2 The vertical-plane-wave model

The vertical-plane-wave model is obtained by assuming barotropic balanced flow, with streamfunction $\psi(x, y, t)$, a uniform background buoyancy frequency N_0 , and a single vertically propagating wave with vertical structure e^{imz} and back-rotated wave velocity $\phi(x, y, t)$. With these idealizations, appendix 4.A derives the vertical-plane-wave model starting from the phase-averaged equations of Wagner and Young [2016]; XV obtain the same model from their version of the phase-averaged equations. In either case, the leading-order wave plus the leading-order balanced velocity (u, v, w) , pressure p , and buoyancy b are:

$$u + iv = e^{i\omega} \phi - \psi_y + i\psi_x; \quad (4.1)$$

$$w = im^{-1} e^{i\omega} \partial \phi + cc; \quad (4.2)$$

$$p = -i\eta e^{i\omega} \partial \phi + cc + f_0 \psi; \quad (4.3)$$

$$b = m\eta e^{i\omega} \partial \phi + cc. \quad (4.4)$$

Above, $\varpi = mz - f_0 t$ is the phase of the vertical plane wave, $\eta = f_0 \lambda^2$ is the wave “dispersivity,” where $\lambda = N_0/f_0 m$ is a horizontal scale, cc denotes complex conjugate, and

$$\partial \stackrel{\text{def}}{=} \frac{1}{2}(\partial_x - i\partial_y) \quad (4.5)$$

is a differential operator. The complex field $\phi(x, y, t)$ in (4.1) is the back-rotated velocity of the near-inertial waves; in (4.2)–(4.4) the other wave fields are expressed in terms of $\partial \phi$. The compact representation of the wave variables in terms of ϕ follows YBJ.

The balanced variables are represented by the streamfunction ψ . Because the Stokes pressure correction is negligible for near-inertial waves [Wagner and Young 2016], the eddies are balanced in the sense that $\psi = \bar{p}/f_0$, where \bar{p} is either the Eulerian-mean or Lagrangian-mean pressure. Moreover, the velocities obtained from ψ in (4.1) are Lagrangian-mean velocities, e.g., the velocity $(-\psi_y, \psi_x)$ advects the material invariant potential vorticity in (4.7) below. We have lightened the notation by using ψ , rather than ψ^L ; the implicit L is particularly important in section 4.6.

The PV of the balanced flow is expressed in terms of ψ and ϕ by

$$q = \underbrace{\Delta \psi}_{\stackrel{\text{def}}{=} \zeta} + \underbrace{\frac{1}{f_0} \left[\frac{1}{4} \Delta |\phi|^2 + \frac{i}{2} J(\phi^*, \phi) \right]}_{\stackrel{\text{def}}{=} q^w}, \quad (4.6)$$

where $\Delta \stackrel{\text{def}}{=} \partial_x^2 + \partial_y^2$ is the horizontal Laplacian and $J(f, g) \stackrel{\text{def}}{=} f_x g_y - f_y g_x$ is the Jacobian, and the superscript star $*$ denotes complex conjugation. Equation (4.6) is the “inversion relation”: q and ϕ determine the Lagrangian-mean flow via $\psi = \Delta^{-1}(q - q^w)$ where q^w defined in (4.6) is the “wave potential vorticity.” Once ψ is obtained by inversion, the field equations (4.7) and (4.8) below can be time-stepped.

Using the generalized Lagrangian-mean formulation Bühler and McIntyre [1998] showed that the assumption of weak interaction between internal waves and balanced flow results in wave-averaged term q^w contributing to the materially conserved PV; see also Grimshaw [1975]. In (4.6) the wave-averaged feedback on the balanced flow is expressed concisely in terms of the backrotated velocity ϕ via the quadratic terms in q^w .

4.2.1 The evolution equations

The balanced flow evolves according to PV advection

$$q_t + J(\psi, q) = D_q; \quad (4.7)$$

the back-rotated velocity satisfies the wave equation

$$\phi_t + J(\psi, \phi) + \phi^{\frac{1}{2}}\zeta - \frac{1}{2}\eta\Delta\phi = D_\phi. \quad (4.8)$$

D_q and D_ϕ in (4.7) and (4.8) are dissipative terms described below.

The wave equation (4.8) is the YBJ model in the case where the near-inertial wave has e^{imz} structure. The back-rotated wave velocity, ϕ , evolves through dispersion—the last term on the left of (4.8)—and nonlinear advection and refraction by the second and third terms in (4.8). Without advection, (4.8) is analogous to Schrödinger’s equation [e.g., Landau and Lifshitz 2013, page 51]. The relative vorticity, $\zeta = \Delta\psi$, is the potential, with negative ζ a well, and the “dispersivity,” $f_0\lambda^2$, is Planck’s constant [Balmforth, Llewellyn Smith, and Young 1998; Balmforth and Young 1999; Danioux, Vanneste, and Bühler 2015]. The quantum analogy may be useful for some readers, but it is not necessary for the understanding of the results below.

The terms on the right of (4.7) and (4.8), D_q and D_ϕ , represent small-scale dissipation. Small-scale dissipation is necessary to absorb the forward transfers of potential enstrophy and wave kinetic and potential energies in the numerical solutions reported below. We find that biharmonic diffusion and viscosity,

$$D_q = -\kappa_e \Delta^2 q \quad \text{and} \quad D_\phi = -\nu_\phi \Delta^2 \phi, \quad (4.9)$$

are sufficient to extend the spectral resolution compared to Laplacian dissipation. In practice, we choose κ_e and ν_ϕ so that the highest 35% of the modes lie in the dissipation range and aliased wavenumbers are strongly damped.

4.2.2 The small-amplitude limit and the validity of the NIW-QG approximation

The development of the NIW-QG model is ordered first by assuming that the waves are weak in the sense that

$$\epsilon \stackrel{\text{def}}{=} \frac{U_w}{f_0 L} \ll 1. \quad (4.10)$$

Above, L is a characteristic scale of both waves and balanced flow and U_w is a characteristic near-inertial wave velocity. The other small parameter in the expansion is the Rossby number of the balanced flow

$$Ro \stackrel{\text{def}}{=} \frac{U_e}{f_0 L} \ll 1, \quad (4.11)$$

where U_e is the eddy velocity. The inequalities in (4.10) and (4.11) must be satisfied in order to obtain the NIW-QG system. But XV and Wagner and Young [2016] make a third assumption: $Ro = \epsilon^2$, or equivalently that $U_e = \epsilon U_w$. The resulting distinguished limit,

$$\epsilon \rightarrow 0, \quad \text{with} \quad Ro = \epsilon^2, \quad (4.12)$$

promotes the importance of wave-averaged effects so that q^w appears at an early, and accessible, order in the expansion. Thus (4.12) is for convenience rather than necessity.

The asymptotic ordering in (4.12) does not imply that the QG–NIW system fails for weaker waves, i.e., if U_w is comparable to, or even smaller than, U_e . Making U_w weaker than U_e delays wave-averaged effects to longer times—it does not, *per se*, invalidate the expansion. The main problem with the weak-wave limit is that other physics, not considered in the NIW-QG system, will contend with wave-averaged effects on ultra-long time scales. For example, even without waves, order Ro^2 ageostrophic effects modify the evolution of balanced flow and produce departures from QG (e.g., see Muraki, Snyder & Rotunno 1999).

To summarize: the main conditions for the validity of the QG–NIW system are (4.10) and (4.11); additionally, validity of the wave equation (4.8) requires that the wave frequency is close of f_0 . The weak-wave limit $U_w/U_e \rightarrow 0$ is valid within the NIW-QG framework: in that limit the system reduces to the barotropic potential vorticity equation and the YBJ equation for a passive wave field.

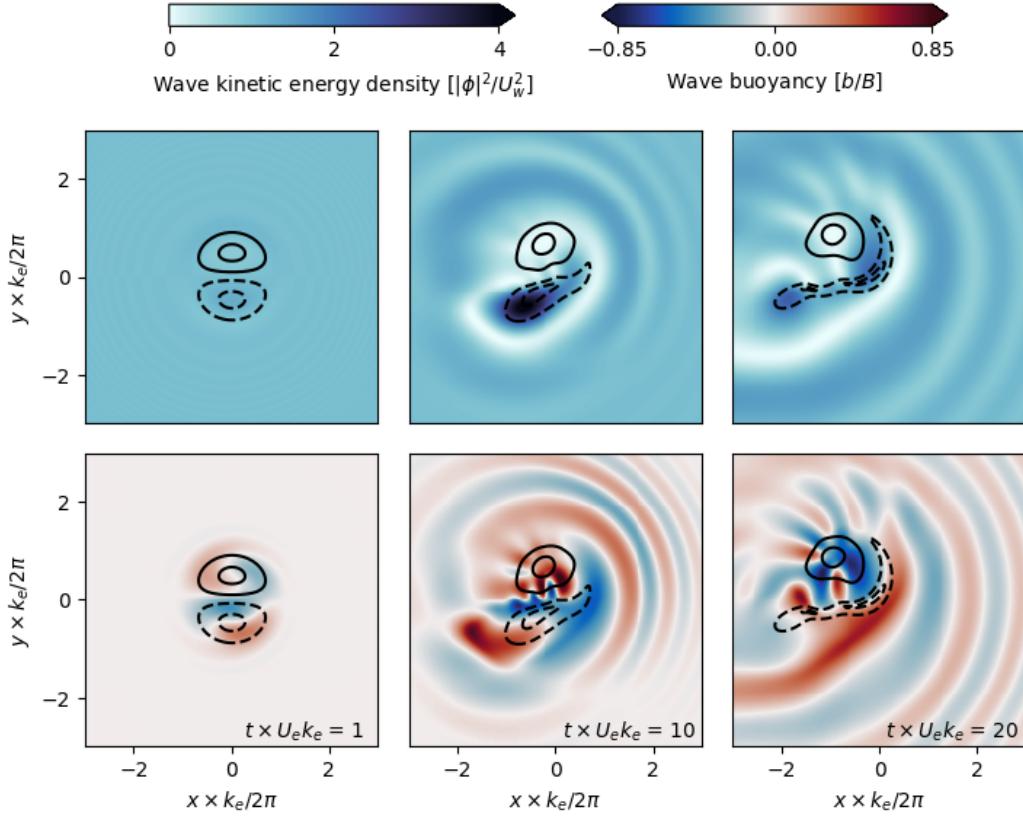


Figure 4.1: Snapshots of the dipole solution with parameters in table 4.2. Contours depict potential vorticity, $q/(U_e k_e) = [-0.5, 0.5, 1.5]$, with dashed lines showing negative values. The upper row shows the wave action density $|\phi|^2/2f_0$. The lower row shows the wave buoyancy; the buoyancy scale is $B = k_e m U_w f_0 \lambda^2$. These plots only show the central $(1/5)^2$ of the simulation domain.

The standard QG approximation is used successfully even when $Ro \sim 1$ [Hoskins 1975] and we expect the NIW-QG model to enjoy similar success if \ll is replaced by $<$ in (4.11). Flows with $Ro > 1$ eddies, such as those reported in Barkan, Winters, and McWilliams [2016], evolve on time scales close to f_0^{-1} , and an Eulerian decomposition into near-inertial waves and eddies is ill-defined unless there is spatial scale separation between eddies and waves. These large Rossby number flows are outside the purview of the NIW-QG model [XV; Wagner and Young 2016].

Table 4.2: Description of parameters of Lamb-Chaplygin simulation. The initial condition have Rossby number $Ro = U_e k_e / f_0 \approx 0.05$, wave dispersivity $\hbar = f_0 \lambda^2 k_e / U_e \approx 1$, and wave amplitude $\alpha = Ro(U_w/U_e)^2 \approx 3.75$.

Parameter	Description	Value
$R = 2\pi k_e^{-1}$	Dipole radius	$L_d/15 \approx 84$ km
U_e	Dipole strength	5×10^{-2} m s $^{-1}$
U_w	NIW speed	5×10^{-1} m s $^{-1}$
N_0	Buoyancy frequency	5×10^{-3} s $^{-1}$
f_0	Coriolis frequency	10^{-4} s $^{-1}$
$2\pi m^{-1}$	NIW vertical wavelength	325 m
κ_e	PVbiharmonic diffusivity	5×10^7 m 4 s $^{-1}$
ν_w	NIW biharmonic viscosity	1×10^7 m 4 s $^{-1}$
N	Number of modes	512
L	Domain size	$2\pi \times 200$ km

4.2.3 An illustrative solution: the Lamb-Chaplygin dipole

As a preamble to our discussion of stimulated generation in freely-evolving 2D turbulence, we consider an example in which the initial QG flow is the Lamb-Chaplygin dipole—see figure 4.1. This dipole is an exact solution of the Euler equations on an infinite two-dimensional plane where the vorticity is confined to a circle of radius R Meleshko and Van Heijst [1994]. The relative vorticity, steady in a frame moving at uniform zonal velocity U_e , is

$$\zeta = \frac{2U_e \kappa}{J_0(\kappa R)} \begin{cases} J_1(\kappa r) \sin \theta, & \text{if } r \leq R, \\ 0, & \text{if } r \geq R. \end{cases} \quad (4.13)$$

Above $r^2 = (x - x_c)^2 + (y - y_c)^2$ is the radial distance from the center (x_c, y_c) , $\tan \theta = (y - y_c)/(x - x_c)$, and J_n is the n 'th order Bessel function. The matching condition at $r = R$ is that $J_1(\kappa R) = 0$ and the smallest solution is $\kappa R \approx 3.8317$. If there is no coupling to the wave ϕ , then the dipole (4.13) is a solution of the QG equation (4.7).

We strongly perturb the dipole in (4.13) by seeding a wave with initial velocity:

$$\phi(x, y, t = 0) = \frac{1+i}{\sqrt{2}} U_w. \quad (4.14)$$

If there was no dipole, this initial condition produces a spatially uniform near-inertial

oscillation with speed U_w . Further parameters of this solution are summarized in table 4.2; note $U_w = 10U_e$.

We integrate the vertical plane-wave model using a standard collocation Fourier spectral method. We evaluate the quadratic non-linearities, including in the wave potential vorticity (4.6), in physical space and transform the product into Fourier space. We time march the spectral equations using an exponential time differencing method with a fourth order Runge-Kutta scheme—for details, see Kassam and Trefethen [2005] and Cox and Matthews [2002].

With the initial condition in (4.14), q^w in (4.6) and $\nabla\phi$ are both zero at $t = 0$. The refractive term in the wave equation (4.8), however, immediately imprints itself onto ϕ thus creating non-zero $\nabla\phi$ and non-zero q^w . Once refraction has created gradients in ϕ , the advective term in (4.8) can further distort ϕ and increase $\nabla\phi$. This scale reduction of ϕ is most evident in the wave buoyancy shown in lower row of figure 4.1. Figure 4.1 also shows the well-known focussing of waves into the negative vortex. But the wave feedback on the mean flow through q^w then results in distortion and shearing of the dipole so that the negative vortex loses its integrity; the lop-sided dipole then starts to drift. Once the negative vortex is distorted to small scales it no longer acts as an effective potential well: the trapping of wave energy by the deformed anti-cyclone is weaker than in the initial condition. In fact, figure 4.1, which shows the materially conserved PV q , understates the development of small scales in the relative vorticity ζ : figure 4.2 shows that both ζ and q^w develop small scales with significant cancellation resulting in the relatively smooth field $q = \zeta + q^w$ shown in figure 4.1. Thus at the final time in figure 4.1 the waves are no longer strongly trapped in the region with $\zeta < 0$. This phenomenology, including significant cancellation between ζ and q^w , is also characteristic of wave-modified two-dimensional turbulence in section 4.4.

To understand the results in figure 4.1 and quantify the stimulated generation of wave energy, we need to understand the conservation laws of the vertical-plane-wave model.

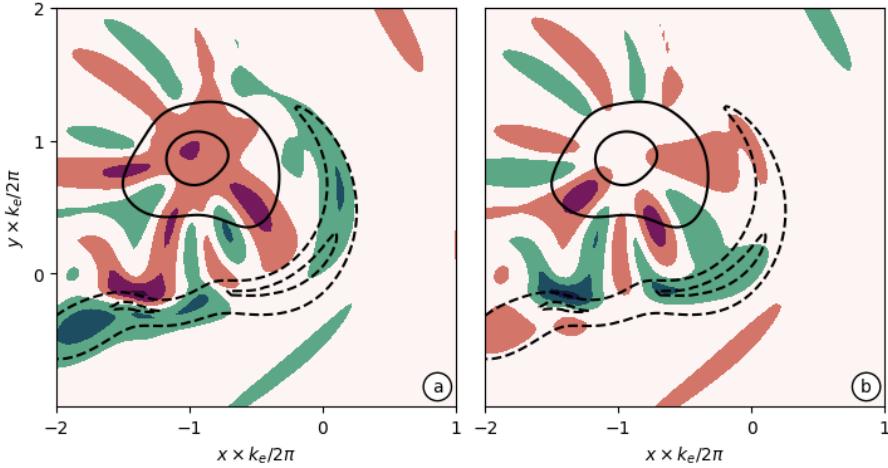


Figure 4.2: (a) Snapshot of q (contours) and ζ (colors) at $t \times U_e k_e = 20$. (b) Snapshot of q (contours) and wave PV q^W (colors). Both black lines and colors depict the contour levels $[-1.5, -0.5, 0.5, 1.5] \times (U_e k_e)$. Solid lines and reddish colors depict positive values; dashed lines and greenish colors show negative values.

4.3 Conservation laws of the vertical-plane-wave model

XV noted that the vertical-plane-wave model in (4.6) through (4.8) inherits two quadratic conservation laws from the parent NIW-QG model: if there is no dissipation then wave action,

$$\mathcal{A} \stackrel{\text{def}}{=} \frac{|\phi|^2}{2f_0}, \quad (4.15)$$

and the energy,

$$\mathcal{E} \stackrel{\text{def}}{=} \frac{1}{2} |\nabla \psi|^2 + \frac{1}{4} \lambda^2 |\nabla \phi|^2, \quad (4.16)$$

are both separately conserved. Following Bretherton and Garrett [1968], the action in (4.15) is the wave energy divided by the intrinsic frequency; YBJ observed that to leading order the wave energy is only kinetic and the intrinsic frequency in (4.15) is the inertial frequency f_0 .

The conserved energy density \mathcal{E} in (4.16) is the sum of the kinetic energy of the balanced flow

$$\mathcal{K} \stackrel{\text{def}}{=} \frac{1}{2} |\nabla \psi|^2, \quad (4.17)$$

and the potential energy of the near-inertial waves,

$$\mathcal{P} \stackrel{\text{def}}{=} \frac{1}{2} b^2 / N_0^2 = \frac{1}{4} \lambda^2 |\nabla \phi|^2. \quad (4.18)$$

Above, $b \propto \partial \phi$ is the wave buoyancy defined in (4.4). We emphasize again that ψ is the streamfunction of the Lagrangian-mean flow, which is in geostrophic balance; the Eulerian-mean flow is not in balance. Thus we refer to \mathcal{K} is the kinetic energy of the balanced flow.

XV explain the physical basis of stimulated generation by noting that balanced kinetic energy \mathcal{K} can be converted into wave potential energy \mathcal{P} while conserving the integral of the total energy \mathcal{E} in (4.16). Indeed, this conversion *must* occur if $\nabla \phi$ is increased by a combination of refraction and advection in the wave equation (4.8). In the example shown in figure 4.1, the initial wave field in (4.14) has infinite spatial scale and therefore there is no wave potential energy at $t = 0$. The subsequent evolution in figure 4.1 involves creation of non-zero $\nabla \phi$, corresponding to gain of \mathcal{P} at the expense of \mathcal{K} : this is stimulated generation of near-inertial waves.

To substantiate this intuition, and diagnose results from our simulation of wave modified two-dimensional turbulence, we develop the conservation laws corresponding to (4.15) and (4.16) in more detail.

4.3.1 Action conservation equation and action flux

Multiplying the wave equation (4.8) by ϕ^* and adding to the complex conjugate we obtain a conservation equation for action density

$$\partial_t \mathcal{A} + J(\psi, \mathcal{A}) + \nabla \cdot \mathcal{F} = \underbrace{\frac{1}{2f_0} (\phi^* D_\phi + \phi D_{\phi^*})}_{\stackrel{\text{def}}{=} D_{\mathcal{A}}}, \quad (4.19)$$

where the flux of near-inertial wave action is

$$\mathcal{F} \stackrel{\text{def}}{=} \frac{i}{4} \lambda^2 (\phi \nabla \phi^* - \phi^* \nabla \phi). \quad (4.20)$$

The local conservation law (5.82) shows how the wave action \mathcal{A} changes due to geostrophic advection and divergence of the wave flux and dissipation—the second, third, and fourth terms in (5.82).

The wave action flux \mathcal{F} is analogous to the probability current of quantum mechanics [e.g., Landau and Lifshitz 2013, pg. 57]. Using the polar representation $\phi = |\phi|e^{i\Theta}$, the wave action flux \mathcal{F} can also be written as

$$\mathcal{F} = \mathcal{A} \eta \nabla \Theta, \quad (4.21)$$

where recall that $\eta = f_0 \lambda^2$ is the dispersivity. In (4.21), $\eta \nabla \Theta$ is the “generalized group velocity” of hydrostatic near-inertial waves, i.e., \mathcal{F} is the generalized group velocity times the action density \mathcal{A} . We use the term “generalized” because no WKB-type scale separation is required to obtain the results above. The connection to standard internal-wave group velocity is quickly verified by considering a plane near-inertial wave with $\Theta = kx + ly$, yielding $\eta \nabla \Theta = N_0^2(k, l)/f_0 m^2$.

Another useful identity involving the action flux \mathcal{F} is

$$\nabla \cdot (\hat{\mathbf{k}} \times \mathcal{F}) = \frac{i}{2} \lambda^2 J(\phi^*, \phi), \quad (4.22)$$

where $\hat{\mathbf{k}}$ is the unit vector perpendicular to the (x, y) -plane. Using (4.22), and the definition of action in (4.15), the wave PV q^w in (4.6) can be written as

$$q^w = \frac{1}{2} \Delta \mathcal{A} + \eta^{-1} \nabla \cdot (\hat{\mathbf{k}} \times \mathcal{F}). \quad (4.23)$$

Denoting an average over the domain by $\langle \rangle$, and assuming that the action flux divergence $\nabla \cdot \mathcal{F}$ vanishes after integration, we obtain from (5.82)

$$\frac{d\langle \mathcal{A} \rangle}{dt} = \varepsilon_{\mathcal{A}}. \quad (4.24)$$

where $\varepsilon_{\mathcal{A}} \stackrel{\text{def}}{=} \langle \phi^* D_\phi + \phi D_{\phi^*} \rangle / (2f_0)$ is the domain average of the dissipative term on the right of (5.82). In the example shown in Figure 4.1 the total action $\langle \mathcal{A} \rangle$ is conserved to within 1% over the course of the integration.

4.3.2 Ehrenfest’s theorem

The quantum analogy suggests that we should seek an analog of Ehrenfest’s theorem (the quantum equivalent of Newton’s law that force equals mass times acceleration). Thus in appendix 4.B we develop a local conservation law for \mathcal{F} . The domain average of that result is

$$\frac{d\langle \mathcal{F} \rangle}{dt} = \hat{\mathbf{k}} \times \langle \zeta \mathcal{F} \rangle - \hat{\mathbf{k}} \times \langle (\mathcal{F} \cdot \nabla) \nabla \psi \rangle - \eta \langle \mathcal{A} \nabla \frac{1}{2} \zeta \rangle + \varepsilon_{\mathcal{F}}. \quad (4.25)$$

In the quantum analogy, \mathcal{F} is momentum and the left hand side of (4.25) is mass times acceleration; the forces are on the right of (4.25). Starting from the end, $\varepsilon_{\mathcal{F}}$ is a dissipative term defined in appendix 4.B. The third term on the right of (4.25) is the force due to the gradient of the potential $\zeta/2$. The second term on the right of (4.25) represents the combined effect of stretching and tilting of \mathcal{F} by the geostrophic flow. The first term on the right of (4.25) is a “vortex force,” again due to ζ , but perpendicular to \mathcal{F} . The two $\hat{\mathbf{k}} \times$ terms on the right lack quantum analogs.

The results in this section are obtained from the wave equation (4.8) without using the PV equation (4.7). In other words, (5.82)-(4.25) apply to the YBJ equation with e^{imz} structure regardless of the balanced-flow dynamics. We turn now to energy conservation and consideration of the PV equation (4.7).

4.3.3 Energy conservation and conversion

The energy conservation law is considerably more complicated than action conservation. We sequester the details of the local conservation laws to appendix 4.B and present here the simpler results obtained by domain averaging those local conservation laws. The results in (4.26) through (4.30) below are obtained by: (1) multiplying the wave equation (4.8) by $\Delta\phi^*$, forming the average $\langle \rangle$, and adding the complex conjugate; and by (2) multiplying the PV equation (4.7) by $-\psi$ and averaging $\langle \rangle$.

For the wave potential energy in (4.18) and the balanced kinetic energy in (4.17) we find

$$\frac{d\langle \mathcal{P} \rangle}{dt} = \Gamma_r + \Gamma_a + \varepsilon_{\mathcal{P}}, \quad (4.26)$$

$$\frac{d\langle \mathcal{K} \rangle}{dt} = -\Gamma_r - \Gamma_a + \Xi + \varepsilon_{\mathcal{K}}, \quad (4.27)$$

where the “conversion terms” in (4.26) and (4.27) are

$$\Gamma_r \stackrel{\text{def}}{=} \left\langle \frac{1}{2} \zeta \nabla \cdot \mathcal{F} \right\rangle, \quad (4.28)$$

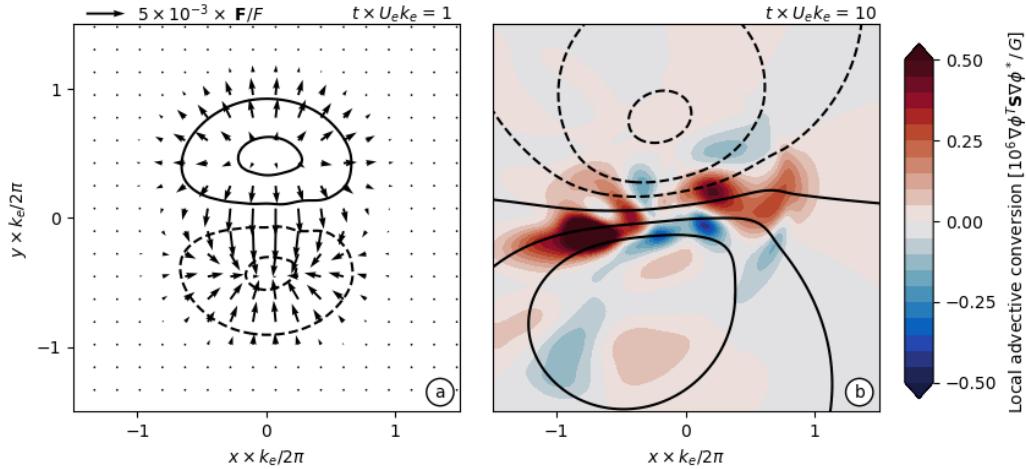


Figure 4.3: Illustration of energy conversion terms in the dipole solution with parameters in table 4.2. (a) The action flux \mathcal{F} overlaid on contours of relative vorticity $\Delta\psi/(U_e k_e) = [-1.5, -0.5, 0.5, 1.5]$, with dashed lines showing negative values; the scale of the action flux is $F = f_0 \lambda^2 k_e U_w^2$. (b) The local advective conversion (i.e., the unaveraged Γ_a) overlaid on contours of streamfunction $\psi \times (U_e/k_e) = [-8, -4, -2, 0, 2, 4, 8]$.

and

$$\Gamma_a \stackrel{\text{def}}{=} \frac{\lambda^2}{4} \langle \Delta\phi^* J(\psi, \phi) + \Delta\phi J(\psi, \phi^*) \rangle, \quad (4.29)$$

$$= -\frac{\lambda^2}{2} \left\langle \begin{bmatrix} \phi_x^* & \phi_y^* \end{bmatrix} \begin{bmatrix} -\psi_{xy} & \frac{1}{2}(\psi_{xx} - \psi_{yy}) \\ \frac{1}{2}(\psi_{xx} - \psi_{yy}) & \psi_{xy} \end{bmatrix} \begin{bmatrix} \phi_x \\ \phi_y \end{bmatrix} \right\rangle. \quad (4.30)$$

The dissipative terms, $\varepsilon_{\mathcal{P}}$, $\varepsilon_{\mathcal{K}}$ and Ξ are defined in appendix 4.B. Ξ in (4.27) is particularly interesting: dissipation of waves D_ϕ produces balanced kinetic energy. Summing (4.26) and (4.27) the “conversion” terms Γ_r and Γ_a cancel, and we obtain the conservation law for total energy $\langle \mathcal{E} \rangle = \langle \mathcal{P} + \mathcal{K} \rangle$.

The refractive conversion Γ_r stems from $i\phi \zeta/2$ in the wave equation and is easy to interpret: the convergence of the wave action flux, $\nabla \cdot \mathcal{F} < 0$, into anti-cyclones, $\zeta < 0$, is a source of wave potential energy \mathcal{P} ; similarly, the divergence of the wave action flux from cyclones is a source of \mathcal{P} . Figure 4.3a shows the convergence of \mathcal{F} into the anti-cyclone (and divergence from the cyclone) of the dipole solution at $t \times U_e k_e = 1$, which yields the sharp initial increase of $\langle \mathcal{P} \rangle$ discussed below. Ehrenfest’s theorem in (4.25) illuminates the initial structure of \mathcal{F} in figure 4.3. Because ϕ is

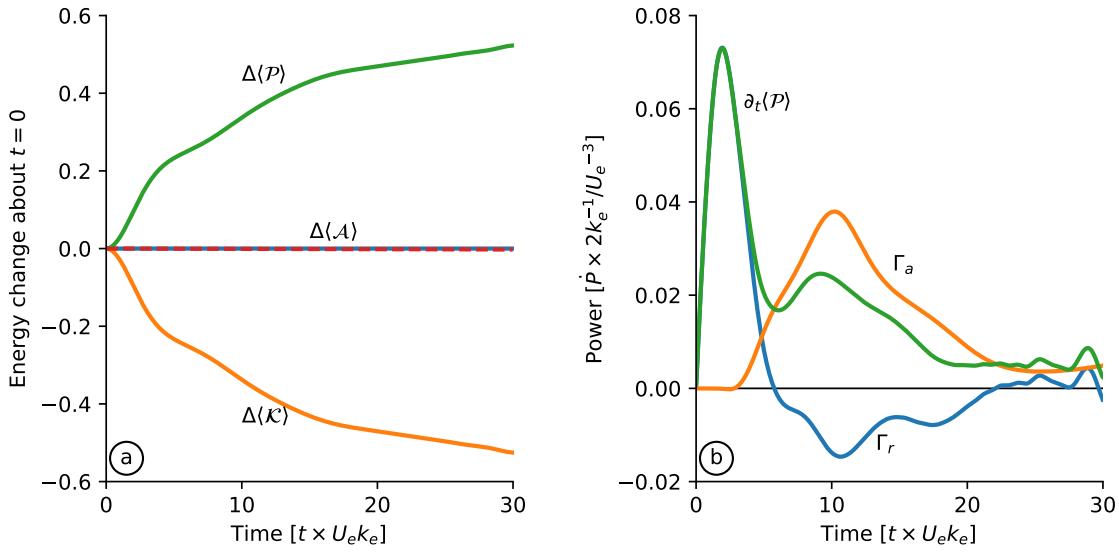


Figure 4.4: Diagnostics of the Lamb-Chaplygin dipole solution with parameters presented in table 4.2. (a) Energy change about initial condition. (b) Wave potential energy budget (4.26).

initially uniform, the initial tendency of \mathcal{P} is

$$\partial_t \mathcal{P}(0) = -\eta \mathcal{A} \nabla_{\frac{1}{2}} \zeta. \quad (4.31)$$

Thus the action flux is initially anti-parallel to the gradient of relative vorticity (see figure 4.3a), and wave action converges into anti-cyclones (and diverges from cyclones).

The advective conversion Γ_a in (4.26) and (4.27) stems from the term $J(\psi, \phi)$ in the wave equation (4.8) and is a source of $\langle \mathcal{P} \rangle$ due to straining and deformation of the wave field by the geostrophic flow. The symmetric 2×2 matrix in (4.30) is the strain or deformation tensor of the geostrophic flow. Thus, in analogy with passive scalar gradient amplification, straining also enhances gradients of the back-rotated near inertial velocity ϕ , thereby generating wave potential energy $\langle \mathcal{P} \rangle$.

4.3.4 Energetics of the Lamb-Chaplygin dipole solution

Figure 4.4 shows the energetics of the Lamb-Chaplygin dipole solution in figure 4.1. In figure 4.4a, \mathcal{P} increases at the expense of \mathcal{K} , while \mathcal{A} is conserved. The wave potential energy budget in figure 4.4b shows that this stimulated generation occurs in two stages. First, refraction of the initially uniform wave field causes a dramatic

Table 4.3: The time-integrated budget of wave potential energy and quasigeostrophic kinetic energy of the Lamb-Chaplygin dipole solutions with parameters provided in table 4.2. The energy budgets close within $10^{-6}\%$.

$\dot{\mathcal{P}}$ budget	Fractional size ($\int \dot{\mathcal{P}} dt / \Delta \mathcal{P}$)	$\dot{\mathcal{K}}$ budget	Fractional size ($\int \dot{\mathcal{K}} dt / \Delta \mathcal{K}$)
Γ_r	0.228	$-\Gamma_r$	-0.227
Γ_a	0.778	$-\Gamma_a$	-0.774
—	—	Ξ_r	0.004
—	—	Ξ_a	0.0
χ_ϕ	-0.006	ϵ_ψ	-0.003
Res.	0.0	Res.	0.0

concentration of waves into the anti-cyclone producing a sharp increase \mathcal{P} through Γ_r . But this rapid initial energy conversion does not last long because the wave feedback deforms the anti-cyclone and dispersion radiates waves away from the dipole (see figure 4.1). Thus in figure 4.4b, Γ_r decreases sharply, and eventually reverses sign at $t \times U_e k_e \approx 8$.

The second stage of stimulated generation starts after refraction has created dipole-scale waves. Advection by the balanced flow can then strain the waves, further reducing their lateral scale (figure 4.4b). The ensuing advective conversion, Γ_a , starts at $t \times U_e k_e \approx 4$. Straining by the balanced flow sustains this advective generation of \mathcal{P} . The waves eventually escape the straining regions through dispersion and the conversion nearly halts at $t \times U_e k_e = 30$. The time-integrated Γ_a accounts for $\approx 78\%$ of the wave potential energy generation; table 4.3 summarizes the energy budget.

4.3.5 Summary

The expressions for energy conversion in (4.28) and (4.30) clarify the mechanism of stimulated generation triggered by the initially uniform near-inertial wave in (4.14). First, refraction causes a convergence of wave action into anti-cyclones. Then advection strains the waves, reducing their lateral scale. Both processes amplify the lateral gradients of wave amplitude, thereby generating wave potential energy \mathcal{P} at the expense of balanced kinetic energy \mathcal{K} . Wave action \mathcal{A} is conserved throughout this process.

In the remainder of this paper, we describe and quantify stimulated generation in an idealization of an oceanographic post-storm scenario: the uniform initial near-inertial wave in (4.14) interacts with two-dimensional turbulence.

4.4 2D turbulence modified by near-inertial waves

To study the energy exchange between near-inertial waves and geostrophic flow in a turbulent regime, we consider a barotropic flow that emerges from random initial conditions integrated for 20 eddy turnover time units. In other words, we first integrate the initial condition

$$\psi(x, y, t \times U_e k_e = -20) = \sum_{k,l} \psi_k \cos(kx + ly + \chi_k) \quad (4.32)$$

with waveless QG dynamics before introducing the wave in (4.14) at $t \times U_e k_e = 0$. Above, χ_k is a random phase uniformly distributed on $[0, 2\pi)$, and ψ_k is the streamfunction isotropic spectrum

$$\psi_k = C \times \{|\mathbf{k}|[1 + (|\mathbf{k}|/k_e)^4]\}^{-1/2}, \quad (4.33)$$

with the wavenumber magnitude $|\mathbf{k}|^2 = k^2 + l^2$. The prescribed initial energy $U_e^2/2$ determines the constant C :

$$\sum_{k,l} \underbrace{|\mathbf{k}|^2 \psi_k^2}_{\stackrel{\text{def}}{=} \mathcal{K}_e} = \frac{1}{2} U_e^2. \quad (4.34)$$

The kinetic energy spectrum, \mathcal{K}_e , peaks at the energy-containing scale k_e^{-1} . At scales larger than k_e^{-1} , \mathcal{K}_e has a linear dependence on $|\mathbf{k}|$, whereas \mathcal{K}_e decays as $|\mathbf{k}|^{-3}$ at scales smaller than k_e^{-1} . This red spectrum ensures insignificant loss of energy by small-scale dissipation D_q in (4.7). Over the course of the simulations described below, the centroid wavenumber of the balanced kinetic energy spectrum decreases by 50%; k_e^{-1} is thus a reasonable scale to characterize the size of the balanced flow throughout the evolution.

In the case with no waves, that is $q^w = 0$, the PV equation (4.7) reduces to 2D fluid mechanics and the quasi-inviscid evolution of a random initial condition is the well-studied problem of 2D turbulence. Stirring of vorticity $\Delta\psi$ transfers enstrophy towards small scales; energy flows to large scales. Most of enstrophy is dissipated

within few eddy turnover times, whereas kinetic energy is nearly conserved. Vorticity concentrates into localized coherent structures: after 20 eddy turnover time units, the vorticity is well-organized into an ensemble of vortices that form via like-sign vortex merging [e.g., Fornberg 1977; McWilliams 1984].

4.4.1 Relevant parameters

The scaling

$$\text{length} \sim k_e^{-1}, \quad \text{time} \sim (U_e k_e)^{-1}, \quad \psi \sim U_e k_e^{-1}, \text{ and} \quad \phi \sim U_w, \quad (4.35)$$

shows that there are two important dimensionless control parameters. The first is

$$\alpha \stackrel{\text{def}}{=} \underbrace{\frac{U_e k_e}{f_0}}_{\stackrel{\text{def}}{=} Ro} \times \left(\frac{U_w}{U_e} \right)^2, \quad (4.36)$$

which scales the contribution of the wave terms in the potential vorticity (4.6). The second dimensionless parameter is

$$\hbar \stackrel{\text{def}}{=} \eta \times \frac{k_e}{U_e} = Ro^{-1} \times (\lambda k_e)^2, \quad (4.37)$$

which scales wave dispersion against the effects of advection and refraction. Assuming that the wave horizontal scale is k_e^{-1} , $(\lambda k_e)^2$ is the wave Burger number, which is small for near-inertial waves.

4.4.2 Solution with $\hbar = 1$ and $\alpha = 0.1$

Figure 4.5 shows snapshots of a solution with $\hbar = 1$ and $\alpha = 0.1$ and further parameters in table 4.4. This turbulence solution shares qualitative aspects of the Lamb-Chaplygin solution. Starting from a uniform wave field in (4.14), refraction quickly concentrates the waves into anti-cyclones. Initially the action density \mathcal{A} is uniform but by $t \times U_e k_e \approx 1$, \mathcal{A} varies on eddy scales by a factor of two with significant focussing of waves, indicated by maxima of \mathcal{A} , into anti-cyclones (compare middle and upper panels of figure 4.5).

Dispersion radiates waves from the vortices; advection enhances the gradients of back-rotated velocity ϕ (see lower panels of figure 4.5, which depict wave buoyancy).

Table 4.4: Description of parameters of the 2D turbulence simulations. The initial condition have Rossby number $Ro = U_e k_e / f_0 \approx 0.05$, wave dispersivity $\kappa = f_0 \lambda^2 k_e / U_e \approx 0.5 - 2$, and wave amplitude $\alpha = Ro(U_w/U_e)^2 \approx 0.1$.

Parameter	Description	Value
$2\pi k_e^{-1}$	Energy-containing scale	$L_d/10 \approx 125 \text{ km}$
U_e	Eddy velocity	$5 \times 10^{-2} \text{ m s}^{-1}$
U_w	NIW speed	$1 \times 10^{-1} \text{ m s}^{-1}$
N_0	Buoyancy frequency	$5 \times 10^{-3} \text{ s}^{-1}$
f_0	Coriolis frequency	10^{-4} s^{-1}
$2\pi m^{-1}$	NIW vertical wavelength	$280 - 560 \text{ m}$
κ_e	PV biharmonic diffusivity	$5 \times 10^6 \text{ m}^4 \text{ s}^{-1}$
ν_w	NIW biharmonic viscosity	$5 \times 10^6 \text{ m}^4 \text{ s}^{-1}$
N	Number of modes	1024
L	Domain size	$2\pi \times 200 \text{ km}$

Table 4.5: The time-integrated budget of wave potential energy and QG kinetic energy of the reference 2D turbulence solution with parameters in 4.4. The energy budgets close within 0.1 %.

\mathcal{P} budget	Fractional size ($\int \mathcal{P} dt / \Delta \mathcal{P}$)	\mathcal{K} budget	Fractional size ($\int \mathcal{K} dt / \Delta \mathcal{P}$)
Γ_r	0.117	$-\Gamma_r$	-0.108
Γ_a	0.907	$-\Gamma_a$	-0.839
—	—	Ξ_r	0.009
—	—	Ξ_a	0.003
χ_ϕ	-0.026	ϵ_ψ	-0.062
Res.	0.003	Res.	0.003

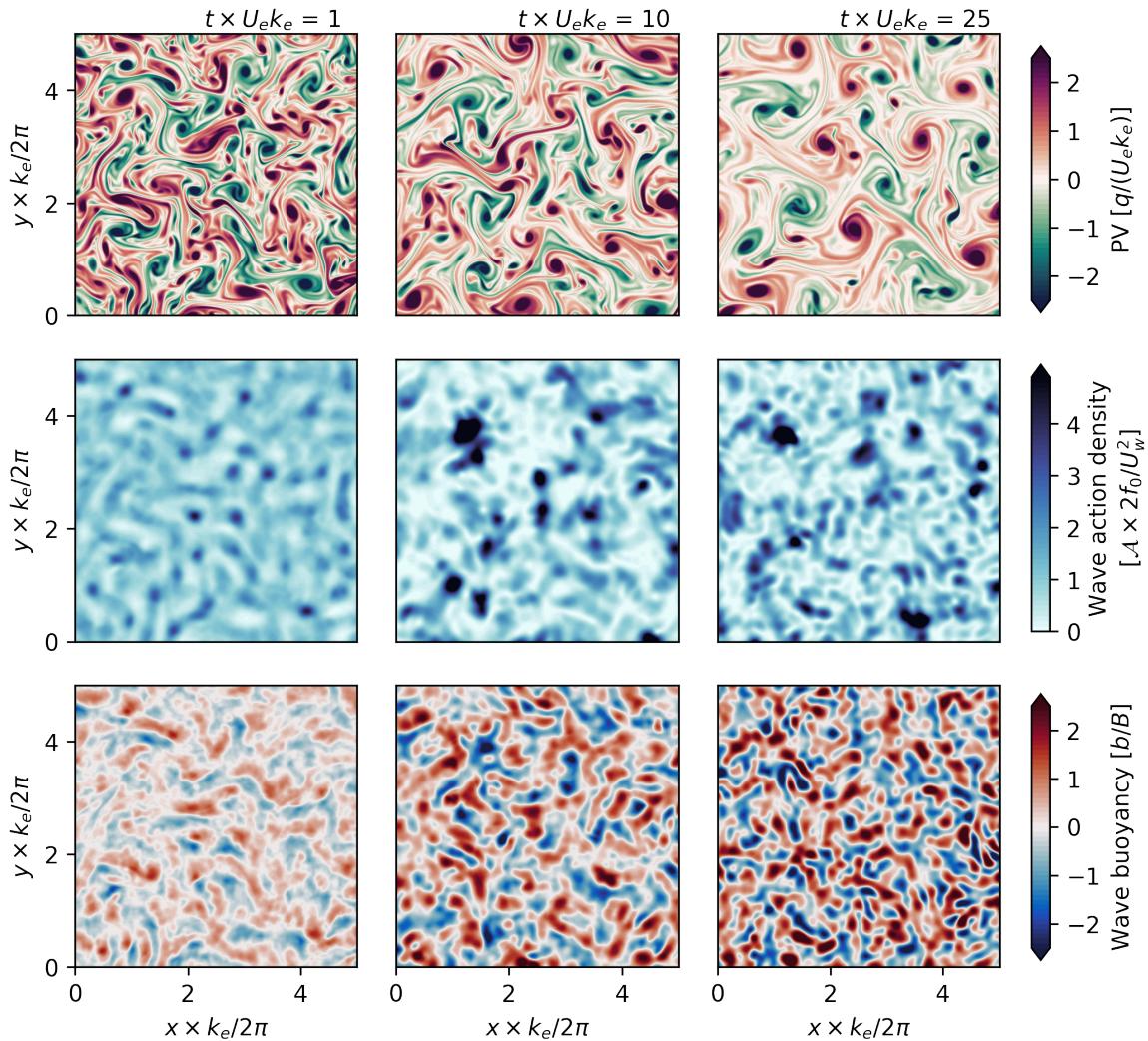


Figure 4.5: Snapshots of the turbulence solution with parameters in table 4.4. Upper panels: PV q . Middle panels: wave kinetic energy density $|\phi|^2$. Bottom panels: wave buoyancy, with scale $B = k_e m U_w f_0 \lambda^2$. These plots show $(1/2)^2$ of the domain.

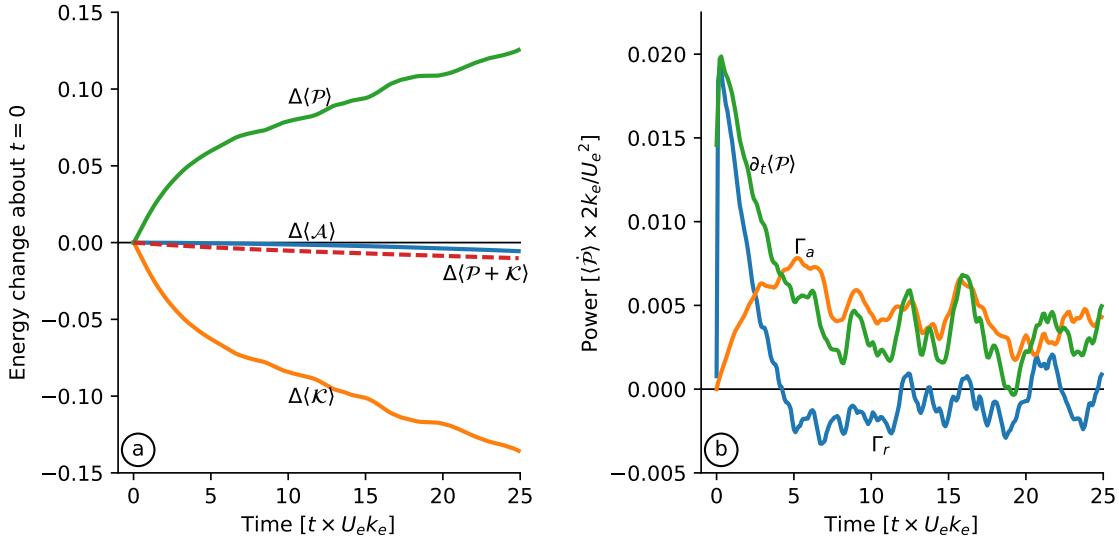


Figure 4.6: Diagnostics of the 2D turbulence solution with parameters presented in table 4.4. (a) Energy change about initial condition. (b) Wave potential energy budget (4.26).

By $t \times U_e k_e \approx 10$, \mathcal{A} varies by a factor of five and the wave buoyancy is amplified by a factor of two. The evolution of potential vorticity q is similar to that in the waveless problem: like-sign vortices merge into bigger vortices. The big vortices keep straining the waves, generating smaller scales in the wave field.

Figure 4.6a shows the inexorable increase in wave potential energy $\langle \mathcal{P} \rangle$ and the corresponding decrease in balanced kinetic energy $\langle \mathcal{K} \rangle$. In figure 4.6b quick wave refraction results in an initial sharp generation of $\langle \mathcal{P} \rangle$ at the expense of balanced kinetic energy $\langle \mathcal{K} \rangle$. As in the Lamb-Chaplygin solution, the positive refractive conversion, $\Gamma_r > 0$, is ephemeral: in figure 4.6b, Γ_r peaks at $t \times U_e k_e \approx 2$ and then decays rapidly, eventually changing sign at $t \times U_e k_e \approx 5$. But a significant positive advective conversion, $\Gamma_a > 0$, sustains stimulated generation so that $\langle \mathcal{P} \rangle$ ultimately increases approximately linearly with time.

After 25 eddy-turnover time units, the balanced kinetic energy $\langle \mathcal{K} \rangle$ has decayed by about 14% from its initial value. Most of this loss is by stimulated generation of $\langle \mathcal{P} \rangle$. As in the Lamb-Chaplygin solution, advective conversion Γ_a accounts for most of the energy change. Table 4.5 presents further details of the energy budget.

The solution illustrates interesting characteristics of stimulated generation. First, the role of refraction is catalytic in that it generates the initial eddy-scale gradients

in ϕ that are then enhanced by advective straining; the advective conversion, Γ_a in (4.30), ultimately accounts for most of the energy transfer from turbulence to waves. Second, the roughly linear-in-time growth of wave potential energy $\langle \mathcal{P} \rangle$ is very slow in comparison with exponential increase of passive-scalar tracer gradients in turbulent velocity fields. The relatively slow growth of $\langle \mathcal{P} \rangle$ suggests that wave dispersion plays an important role in slowing and perhaps opposing advective straining (see section 4.5 for further discussion of dispersion and “wave escape”). To investigate whether these characteristics are general we consider solutions with varying vertical wavelengths and therefore different dispersivities.

4.4.3 Varying dispersivity

Figure 4.7 shows snapshots of potential vorticity q and its constituents in a set of solutions with varying the vertical wavelength $2\pi m^{-1}$ from 280 to 560 m, yielding dispersivities ranging from 0.5 to 2. (All other parameters are fixed.) The potential vorticity q shows more small-scale filamentation with decreasing dispersivity, but it is otherwise similar across the three solutions. The partition into relative vorticity $\Delta\psi$ and wave potential vorticity q^w , however, depends significantly on dispersivity. In particular, q^w develops smaller scales and larger amplitudes with decreasing dispersivity. As anticipated by the dipole example in figure 4.2, there is cancellation of small-scale features in q^w against those in ζ so that q is relatively smooth even in the solution with weak dispersion $\hbar = 0.5$.

The initial evolution of the uniform wave field is similar across dispersivities, with refraction initially generating eddy-scale gradients of the waves—see figure 4.8. Refraction produces a sharp initial increase of wave potential energy and decrease of balanced kinetic energy, which is almost independent of dispersivity. Figure 4.8c shows that this initial “refractive stage” yields a strongly negative wave-vorticity correlation r ,

$$r \stackrel{\text{def}}{=} \frac{\langle \zeta \mathcal{A}' \rangle}{\sqrt{\langle \zeta^2 \rangle \langle \mathcal{A}'^2 \rangle}}, \quad (4.38)$$

where $\mathcal{A}' \stackrel{\text{def}}{=} (|\phi|^2 - |\langle \phi \rangle|^2)/f_0$; in figure 4.8c the early negative r is nearly independent of dispersivity. Because significant energy exchange takes place in the anti-cyclones

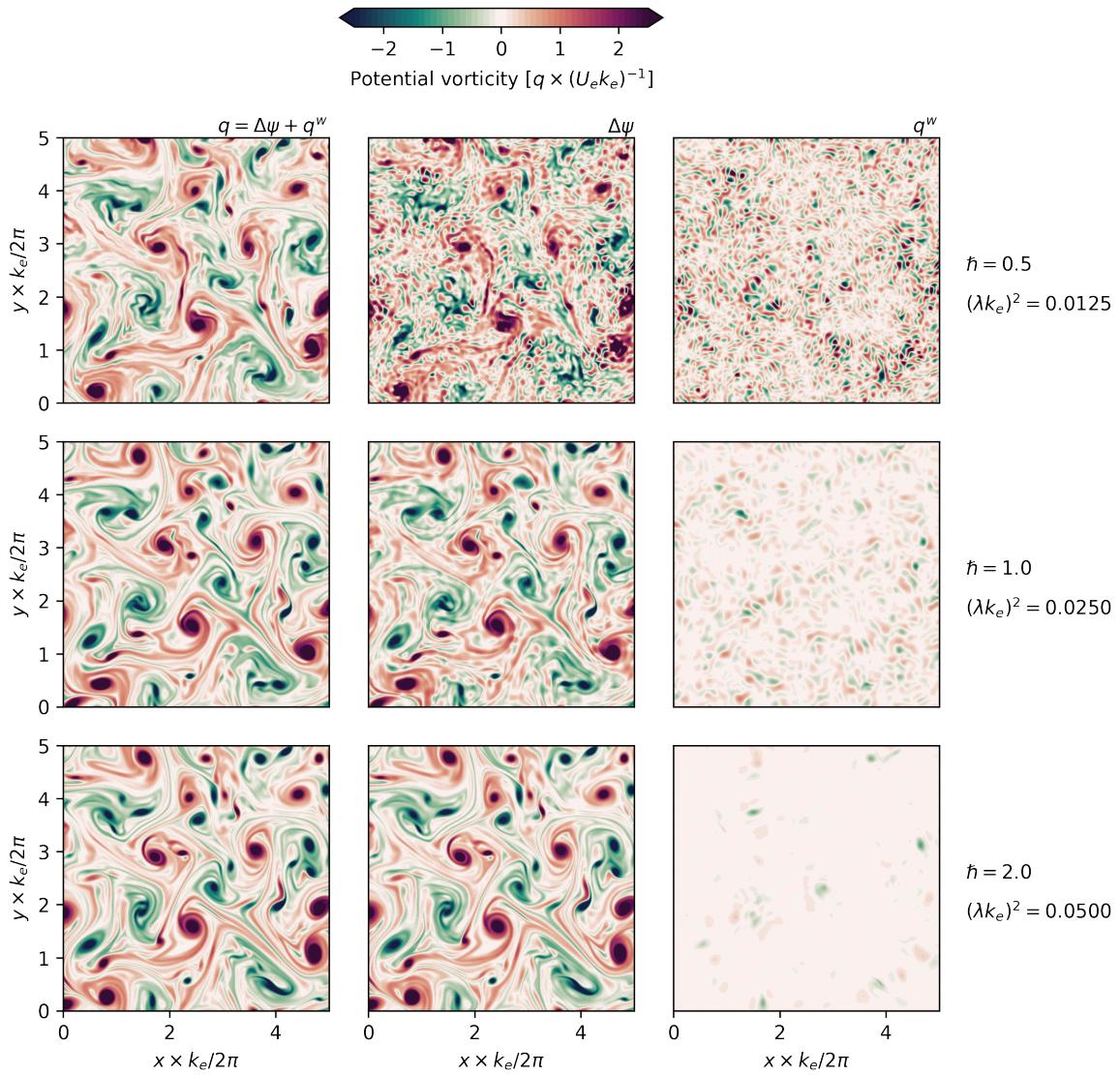


Figure 4.7: Snapshots of PV q and its decomposition into relative vorticity $\zeta = \Delta\psi$, and wave potential vorticity q^w . The snapshots were taken at $t \times U_e k_e = 25$.

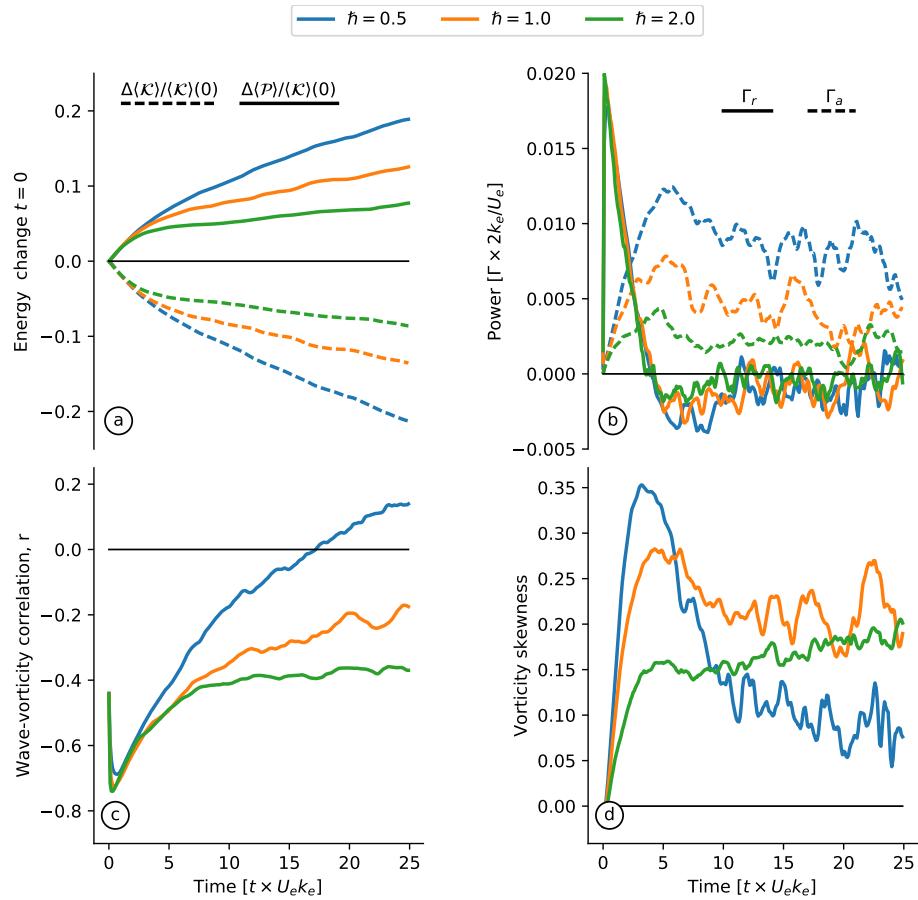


Figure 4.8: The energetics of 2D turbulence solutions with different dispersivities. (a) Energy change about the initial condition. (b) The energy conversion terms in (4.26). (c) The correlation between relative vorticity and wave kinetic energy. (d) The skewness of relative vorticity.

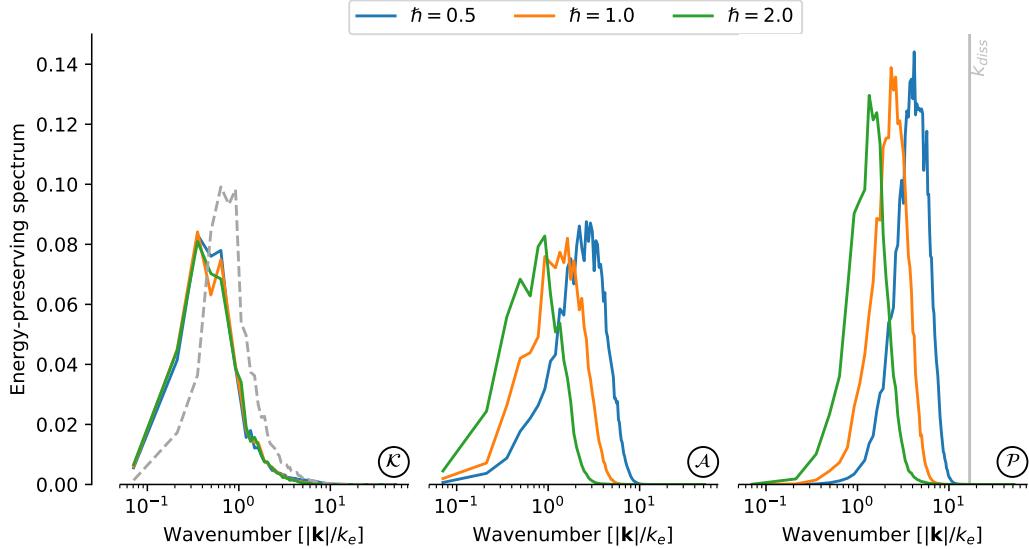


Figure 4.9: Energy-preserving spectra of 2D turbulence solutions with different dispersivities. The three panels show spectra of balanced kinetic energy \mathcal{K} , wave action \mathcal{A} , and wave potential energy \mathcal{P} . All solid lines correspond to spectra at $t \times U_e k_e = 25$ and the dashed line in \mathcal{K} is the balanced kinetic energy spectrum at $t \times U_e k_e = 0$. All spectra are normalized by their total energy, i.e., the area under each curve is one.

due to the initial wave concentration, a positive vorticity skewness ensues (figure 4.8d). Once the eddy-scales are created, advection strains the waves and generates further wave potential energy at the expense of balanced kinetic energy. It is in this stage that the dependence on dispersivity is pronounced: weakly dispersive waves are strained further than strongly dispersive waves. Thus the advective conversion becomes stronger with decreasing dispersivity (figure 4.8b). Advection and dispersion significantly reduces the wave-vorticity correlation; the reduction in correlation increases as the dispersivity decreases (figure 4.8c). For the weakest dispersivity considered, the wave-vorticity correlation becomes weakly positive likely because of the early positive vorticity skewness.

In all solutions reported above, the evolution of the balanced flow is similar to that of waveless 2D turbulence: there is a transfer of balanced energy towards larger scales driven by merger of like-signed vortices—see the left panel of figure 4.9. The main difference is that balanced kinetic energy is constantly transformed into wave potential energy via stimulated generation. The stimulated generation process is

associated with a forward transfer of wave action \mathcal{A} from the infinite horizontal scale in the initial condition (4.14) to the eddy scale; see the middle panel of figure 4.9. The wave potential energy density \mathcal{P} on the right panel of figure 4.9 develops significantly smaller scales than those of the balanced kinetic energy \mathcal{K} .

4.5 Wave escape

The expression for the energy conversion in (4.26) illuminates the physics of stimulated generation: both convergence of wave action density into anti-cyclones (4.28) and geostrophic straining of the wave field (4.30) are sources of wave potential energy and sinks of balanced kinetic energy. But this characterization of stimulated generation ignores the important role of wave dispersion—waves can propagate out of the vorticity or straining regions, thereby reducing the correlations Γ_r and Γ_a required for stimulated generation.

Wave dispersion is the only mechanism that upsets stimulated generation in the quasi-inviscid solutions described in this paper. In all solutions, after an initial conversion due to refraction, advective straining accounts for most of the energy conversion. Experience with the passive-scalar problem suggests (incorrectly) that the wave potential energy \mathcal{P} should then increase exponentially with time as $\nabla\phi$ is amplified by stirring [e.g., Young, Rhines, and Garrett 1982]. But even in the weakly dispersive limit, the waves do not behave as a passive scalar and stimulated generation is much less effective than suggested by this “passive-scalar thinking.” This is because advective straining can only increase $\nabla\phi$ so much: the near-inertial generalized group velocity is $\eta\nabla\Theta$ (cf. section 4.3), where Θ is the phase of the near-inertial back-rotated velocity $\phi = |\phi|e^{i\Theta}$. Geostrophic straining enhances $\nabla\Theta$ thereby increasing the near-inertial group velocity so that the waves escape the straining region. Thus, straining by a barotropic balanced flow results in near-inertial “wave escape,” as opposed to the “wave capture” described by Bühler and McIntyre [2005]. Indeed, wave capture requires *both* lateral strain *and* vertical shear: the vertical-plane-wave model has no vertical shear and therefore wave capture is inoperative; see Thomas [2012] for further discussion of the importance of vertical shear to wave capture.

4.5.1 Strain flow

We are surprised by the successful resistance mounted by the waves to strain-driven exponential amplification of $\nabla\phi$ and thus seek to illustrate wave escape with simple flows. We first consider the straining flow $\psi = -\alpha xy$. Ignoring dissipation for simplicity, the wave equation (4.8) reduces to

$$\phi_t + \alpha x \phi_x - \alpha y \phi_y - \frac{i}{2} \eta \Delta \phi = 0, \quad (4.39)$$

and the action equation is

$$\mathcal{A}_t + \alpha(x\mathcal{A})_x - \alpha(y\mathcal{A})_y + \nabla \cdot \mathcal{F} = 0. \quad (4.40)$$

Without vorticity and dissipation, Ehrenfest's theorem (4.25) reduces to

$$\partial_t [\langle \mathcal{F}^x \rangle, \langle \mathcal{F}^y \rangle] = [-\alpha \langle \mathcal{F}^x \rangle, +\alpha \langle \mathcal{F}^y \rangle], \quad (4.41)$$

with solution

$$[\langle \mathcal{F}^x \rangle, \langle \mathcal{F}^y \rangle] = [\langle \mathcal{F}^x \rangle_0 e^{-\alpha t}, \langle \mathcal{F}^y \rangle_0 e^{\alpha t}], \quad (4.42)$$

where the subscript 0 denotes the initial condition. For a compact wave packet, with a well-defined uniform group velocity c_g , the action flux is

$$\langle \mathcal{F} \rangle = c_g \langle \mathcal{A} \rangle. \quad (4.43)$$

Because $\langle \mathcal{A} \rangle$ is constant, (4.42) is thus also a solution for $c_g(t)$.

Figure 4.10: The escape of a Gaussian near-inertial wave packet from the saddle of a strain flow. The Gaussian decay scale is $\mu = \nu$. The waves are weakly dispersive: $\hbar = \eta/\alpha\mu^2 \approx 0.1$. Black contours show the streamfunction $\psi = -\alpha xy$, and colors represent wave zonal velocity; the colorbar limits are fixed. The black dot indicates the center of the packet and the gray line tracks its trail.

The position of the center of the packet is

$$\mathbf{x}_c(t) \stackrel{\text{def}}{=} \langle \mathbf{x} \mathcal{A} \rangle / \langle \mathcal{A} \rangle. \quad (4.44)$$

Multiplying (4.40) by \mathbf{x} , averaging, and using (4.42), yields

$$\partial_t x_c - \alpha x_c = c_g^x e^{-\alpha t}, \quad \text{and} \quad \partial_t y_c + \alpha y_c = c_g^y e^{\alpha t}, \quad (4.45)$$

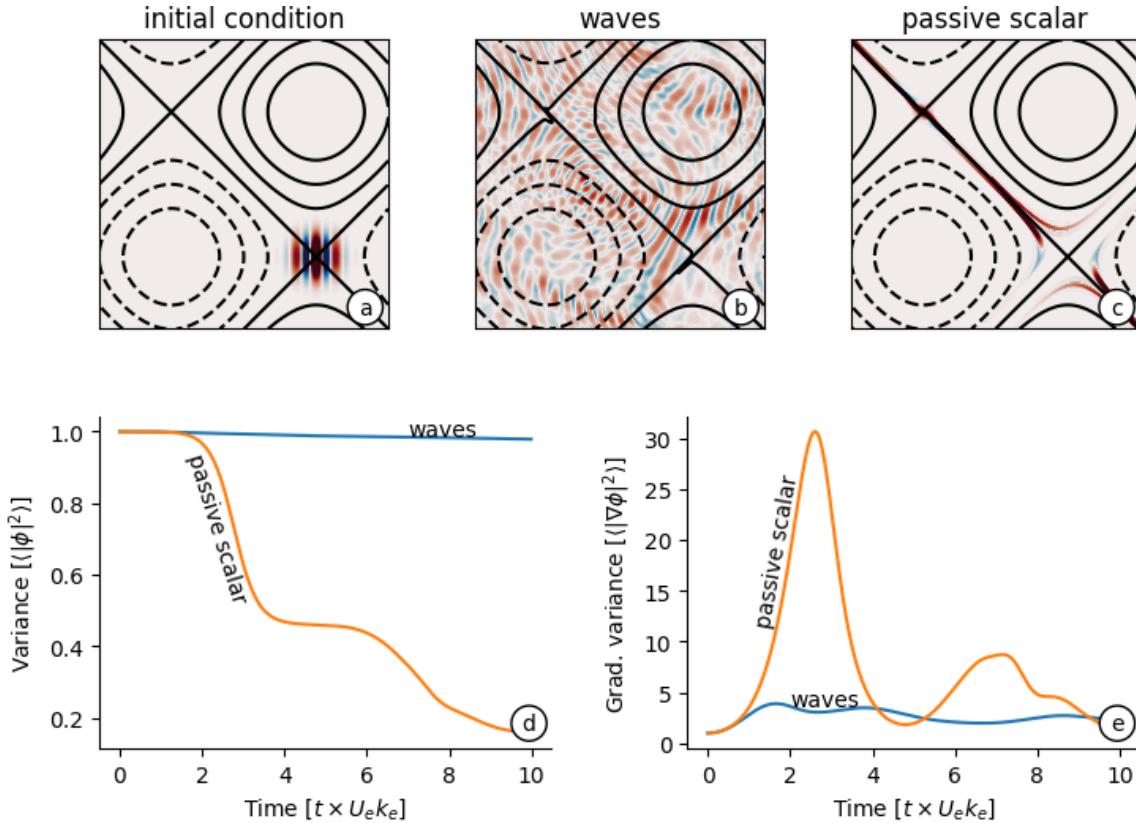


Figure 4.11: A comparison of passive-scalar and wave solutions ($\hbar \approx 0.09$) with same initial conditions (the wave kinetic energy is equal to the passive-scalar variance) and the same small-scale dissipation. (a) Initial condition of wave back-rotated zonal velocity and of the passive scalar. (b) Wave back-rotated zonal velocity at $t \times U_e k_e = 10$. (c) Passive-scalar concentration at $t \times U_e k_e = 10$. (d) Variance of wave velocity or passive-scalar variance. (e) Variance of wave velocity gradient or passive scalar. In (d) and (e), the diagnostics are normalized by their initial values.

with solution

$$x_c(t) = x_{c0} + c_{g0} \frac{\sinh \alpha t}{\alpha}. \quad (4.46)$$

The trajectory of the packet is therefore a straight line following the initial group velocity c_{g0} . The wave packet is not deflected by the hyperbolic streamlines of the straining flow and the packet escapes by accelerating exponentially with time along a straight-line trajectory. This requires that the group velocity $c_g(t)$ adjusts in magnitude and direction so as to keep the center of the packet on the straight and narrow.

The result in (4.46) is so remarkable that it is reassuring to obtain it without invoking Ehrenfest's theorem. We thus consider the specific example of a wave packet

launched in the strain field with the Gaussian initial condition

$$\phi(x, y, 0) = \frac{2\pi}{\mu\nu} \exp \left[-\frac{1}{2} \left(\frac{x^2}{\mu^2} + \frac{y^2}{\nu^2} \right) + i(px + qy) \right]. \quad (4.47)$$

If $p\mu \gg 1$ and $q\nu \gg 1$ this is a wave packet with initial wavenumber (p, q) and initial group velocity

$$\mathbf{c}_{g0} = \eta(p, q). \quad (4.48)$$

The exact solution to (4.39), subject to the initial condition (4.47), is

$$\begin{aligned} \phi = & \frac{2\pi}{\sqrt{(\mu^2 + i\eta f)(\nu^2 + i\eta g)}} \times \\ & \exp \left[-\frac{1}{2} \frac{x^2 e^{-2\alpha t} - 2i\mu^2 p x e^{-\alpha t} + i\eta f \mu^2 p^2}{\mu^2 + i\eta f} - \frac{1}{2} \frac{y^2 e^{2\alpha t} - 2i\nu^2 q y e^{\alpha t} + i\eta g \nu^2 q^2}{\nu^2 + i\eta g} \right], \end{aligned} \quad (4.49)$$

with

$$f \stackrel{\text{def}}{=} \frac{1 - e^{-2\alpha t}}{2\alpha} \quad \text{and} \quad g \stackrel{\text{def}}{=} \frac{e^{2\alpha t} - 1}{2\alpha}. \quad (4.50)$$

Figure 4.10 illustrates the solution (4.49). The strain flow tilts the packet to align it with the x -axis. The flow then exponentially stretches the packet, which in turn escapes from the saddle point. To calculate the trajectory of the packet we note that

$$|\phi|^2 = \frac{(2\pi)^2}{\sqrt{(\mu^4 + \eta^2 f^2)(\nu^4 + \eta^2 g^2)}} \exp \left[-\frac{\mu^2(x e^{-\alpha t} - \eta f p)^2}{\mu^4 + \eta^2 f^2} - \frac{\nu^2(y e^{\alpha t} - \eta g q)^2}{\nu^4 + \eta^2 g^2} \right], \quad (4.51)$$

and we recover $\mathbf{x}_c(t)$ in (4.46) as the center of the packet in (4.51).

4.5.2 A flow with strain and vorticity

As a second example of wave escape, figure 4.11 shows a numerical solution for a wave packet launched at the saddle point of a large-scale balanced flow with $\psi = \sin x + \sin y$. The small-time evolution of the packet is predicted by the strain-flow solution discussed above. The top row of figure 4.11 shows that the behavior of the wave packet is qualitatively different from that of a passive scalar in the same flow: the passive scalar is stretched out along the separatrices while the wave packet escapes into the vortex centers. Thus the passive scalar packet is strained and quickly diffused

into oblivion. On the other hand, the waves are strained just so much, resulting in acceleration and escape from the straining region; the waves finally concentrate in the regions with non-zero vorticity, i.e., in the regions where the Okubo-Weiss criterion indicates no exponential stretching.

On the bottom row, figure 4.11d shows that while wave action $\propto |\phi|^2$ is nearly conserved, the analogous passive-scalar variance is strongly dissipated. Figure 4.11e shows that the variance of the passive-scalar gradient at first increases exponentially due to straining and then decays due to diffusion. On the other hand, the potential energy of the waves $\propto |\nabla\phi|^2$ increases slowly and then oscillates around an equilibrium level. The wave-escape phenomenology in the turbulence solutions of section 4.4 qualitatively resembles that seen in this simple flow. In particular, the wave potential energy does not reach the dissipative scale (figure 4.9c).

4.6 The Eulerian-mean viewpoint

In section 4.3 we developed conservation laws for \mathcal{A} , \mathcal{P} and \mathcal{K} based on the Lagrangian-mean streamfunction ψ and the backrotated velocity ϕ . Material conservation of PV is central to the NIW-QG system and thus the *Lagrangian-mean velocity*, $(-\psi_y, \psi_x)$, must be a key field. But then $\mathcal{K} = |\nabla\psi|^2/2$ is not the kinetic energy of the *Eulerian-mean velocity* and therefore Reynolds stress diagnostics do not directly indicate changes in \mathcal{K} . In fact Reynolds stresses have not been mentioned in section 4.3 and one might wonder how these standard statistics are related to the NIW-QG energy conversions Γ_r and Γ_a ? In this section we identify the Eulerian-mean energy transfers in terms of NIW-QG fields.

4.6.1 Kinetic energy of the Lagrangian-mean and Eulerian-mean flows

appendix 4.A shows that in the vertical-plane-wave model, the Stokes drift is horizontally non-divergent, with streamfunction $-\mathcal{A}$. Thus the Stokes velocity is

$$\mathbf{u}^S = -\hat{\mathbf{k}} \times \nabla \mathcal{A}, \quad (4.52)$$

and the Eulerian-mean streamfunction is

$$\psi^E \stackrel{\text{def}}{=} \psi + \mathcal{A}. \quad (4.53)$$

To illustrate the important differences between Lagrangian-mean and Eulerian-mean viewpoints, figure 4.12 shows a snapshot of the dipole example from section 4.2.3. This snapshot was taken at $t \times U_e k_e = 10$, just after the end of the refractive stage of energy conversion, when waves are strongly concentrated in the anti-cyclone. While ψ is fairly symmetric, the Eulerian-mean streamfunction ψ^E displays a stronger anti-cyclone. The Stokes drift is concentrated in the negative ζ region and is anti-parallel to the Eulerian-mean flow. Thus the asymmetry in ψ^E is compensated by a strong “Stokes cyclone,” which is set up during the refractive stage and thus the Lagrangian-mean streamfunction ψ is more nearly a symmetric dipole.

Using the decomposition (4.53), the balanced kinetic energy $\mathcal{K} = |\nabla\psi|^2/2$ is

$$\mathcal{K} = \underbrace{\frac{1}{2}|\nabla\psi^E|^2}_{\stackrel{\text{def}}{=} \mathcal{K}^E} + \underbrace{\frac{1}{2}|\nabla\mathcal{A}|^2 - \nabla\psi^E \cdot \nabla\mathcal{A}}_{\stackrel{\text{def}}{=} \mathcal{K}^S}. \quad (4.54)$$

\mathcal{K} does not diagonalize, i.e., the cross-term $\nabla\psi^E \cdot \nabla\mathcal{A}$ is non-zero, and the “Stokes kinetic energy” \mathcal{K}^S is not sign definite.

The first goal here is to obtain an expressions for the time rate of change of \mathcal{K}^S and \mathcal{K}^E . We begin with yet another \mathcal{A} - \mathcal{F} identity: using the definitions of \mathcal{A} and \mathcal{F} in (4.15) and (4.20) yields

$$\mathcal{A} \nabla \cdot (\hat{k} \times \mathcal{F}) = -\mathcal{F} \cdot \hat{k} \times \nabla \cdot \mathcal{A}. \quad (4.55)$$

The average of (4.55), combined with a standard vector identities, results in

$$\langle \mathcal{A} \nabla \cdot (\hat{k} \times \mathcal{F}) \rangle = \langle \mathcal{F} \cdot \hat{k} \times \nabla \cdot \mathcal{A} \rangle = 0. \quad (4.56)$$

Equation (4.56) implies that the Stokes velocity u^S is, on average, orthogonal to the action flux \mathcal{F} .

Forming $\langle \mathcal{A} q \rangle$, and combining the expression for q^W in (4.23) with the identity (4.56), yields

$$\langle \mathcal{K}^S \rangle = \langle \mathcal{A} q \rangle. \quad (4.57)$$

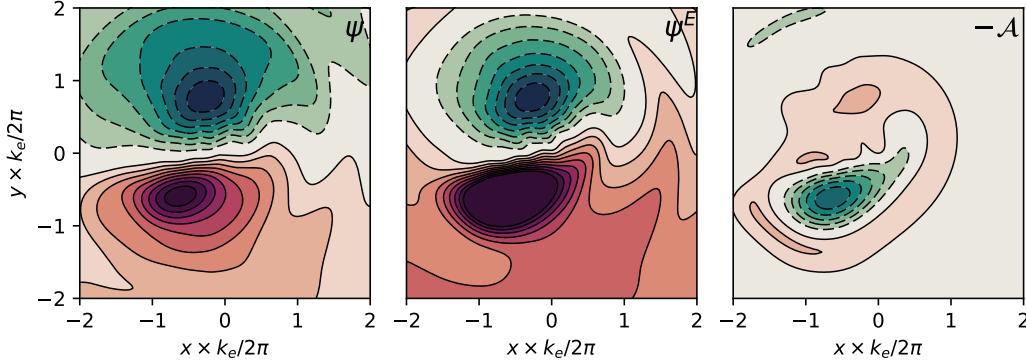


Figure 4.12: Snapshot of the dipole Lagrangian-mean streamfunction and its decomposition into Eulerian-mean and Stokes drift. The snapshots is at $t \times U_e k_e = 10$, just after the end of the refractive stage. The spatial average of each field was removed for direct comparison.

An expression for the rate of change of $\langle \mathcal{K}^S \rangle$ follows by combining the PV advection equation (4.7) with the action conservation (5.82):

$$\frac{d\langle \mathcal{K}^S \rangle}{dt} = -\Gamma_S + \underbrace{\langle \mathcal{A} D_q \rangle + \langle q D_{\mathcal{A}} \rangle}_{\stackrel{\text{def}}{=} \varepsilon_S}, \quad (4.58)$$

where

$$\Gamma_S \stackrel{\text{def}}{=} \langle q \nabla \cdot \mathcal{F} \rangle. \quad (4.59)$$

Finally we obtain the rate of change of the Eulerian-mean kinetic energy by combining (4.58) with (4.27):

$$\frac{d\langle \mathcal{K}^E \rangle}{dt} = \Gamma_S - \Gamma_r - \Gamma_a + \Xi + \varepsilon_{\mathcal{K}} - \varepsilon_S. \quad (4.60)$$

4.6.2 Reynolds stresses and buoyancy fluxes

Within the Eulerian framework the horizontal velocity is represented as

$$\mathbf{u} = \mathbf{u}^E + \tilde{\mathbf{u}}, \quad (4.61)$$

where superscript E denotes the Eulerian-mean and tilde denotes the near-inertial wave velocity. With this decomposition, the Eulerian-mean velocity satisfies

$$\mathbf{u}_t^E + \mathbf{u}^E \cdot \nabla \mathbf{u}^E + (\tilde{\mathbf{u}} \cdot \nabla \tilde{\mathbf{u}})^E + \hat{\mathbf{k}} \times f_0 \mathbf{u}^E + \nabla p^E = 0. \quad (4.62)$$

(Dissipation is neglected in the following discussion.) As in the vertical plane-wave model, we confine attention to mean velocities independent of z , so that $\mathbf{u}^E = (-\psi_y^E, \psi_x^E)$

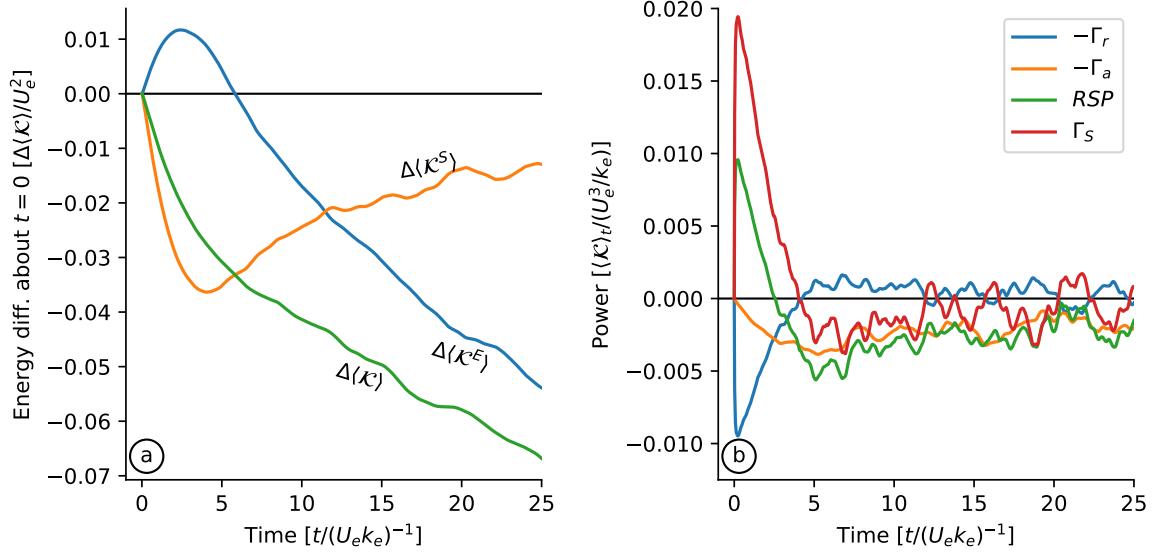


Figure 4.13: Diagnostics of the 2D turbulence solution with parameters presented in table 4.4. (a) Kinetic energy change about initial condition. (b) Refractive and advective conversion terms, and shear production.

where ψ^E is the Eulerian-mean streamfunction introduced in (4.53). Forming $\langle \mathbf{u}^E \cdot (4.62) \rangle$ yields

$$\frac{d\langle \mathcal{K}^E \rangle}{dt} = \underbrace{\left\langle (\tilde{v}^2 - \tilde{u}^2)^E \psi_{xy}^E + (\tilde{u}\tilde{v})^E (\psi_{xx}^E - \psi_{yy}^E) \right\rangle}_{\stackrel{\text{def}}{=} RSP}, \quad (4.63)$$

Above RSP is the Reynolds shear production of Eulerian–mean kinetic energy. Comparing (4.63) with (4.60), and ignoring the dissipative contributions, we see that

$$RSP = \Gamma_s - \Gamma_r - \Gamma_a. \quad (4.64)$$

We can also consider the source of near-inertial potential energy by starting with the wave buoyancy equation

$$\tilde{b}_t + \mathbf{u}^E \cdot \nabla \tilde{b} + \nabla \cdot [\tilde{u}\tilde{b} - (\tilde{u}\tilde{b})^E] + \tilde{w}N^2 = 0, \quad (4.65)$$

where we have confined attention to the vertical-plane-wave model so that $b^E = 0$. Multiplying (4.65) by \tilde{b} , taking the Eulerian average, and then the domain average $\langle \rangle$,

we have

$$\frac{d\langle \mathcal{P} \rangle}{dt} = -\underbrace{\langle (\tilde{w}\tilde{b})^E \rangle}_{\stackrel{\text{def}}{=} BF}, \quad (4.66)$$

$$= \Gamma_r + \Gamma_a. \quad (4.67)$$

Above $\mathcal{P} = (\tilde{b}^2)^E / 2N^2$ is the wave potential energy and BF is the buoyancy flux; we have used (4.26) to relate Γ_r and Γ_a to the Eulerian-mean buoyancy flux.

Finally, from balanced energy equation (4.27) and $\mathcal{K} = \mathcal{K}^E + \mathcal{K}^S$, we deduce that

$$\frac{d\langle \mathcal{K}^S \rangle}{dt} = BF - RSP, \quad (4.68)$$

$$= -\Gamma_S. \quad (4.69)$$

Figure 4.13a shows the decomposition of $\langle \mathcal{K} \rangle$ into $\langle \mathcal{K}^E \rangle$ and $\langle \mathcal{K}^S \rangle$ for the main 2D turbulence solution discussed in section 4.4. In the initial condition there is no Stokes flow and $\mathcal{K} = \mathcal{K}^E$. Then in the refractive stage, the kinetic energy of the Eulerian-mean flow $\langle \mathcal{K}^E \rangle$ increases, while kinetic energy of the Lagrangian-mean flow $\langle \mathcal{K} \rangle$ decreases; the “Stokes energy”, $\mathcal{K}^S = \mathcal{K} - \mathcal{K}^E$, is initially zero and becomes negative, due to a positive ψ^E - \mathcal{A} correlation (cf. figure 4.12). As illustrated in figure 4.12, the refractive stage, which creates the strong spatial modulations in action density \mathcal{A} , can also be viewed as the setting-up the Stokes velocity in (4.52).

In the refractive stage, the advective conversion Γ_a is small and $\Gamma_S \approx 2\Gamma_r$. Hence the shear production is due to the refractive conversion, $RSP \approx \Gamma_r$ (see figure 4.13b). At later times, during the advective stage, $\Gamma_S - \Gamma_r$ is small, and therefore $RSP \approx BF \approx -\Gamma_a$. At this stage, $\partial_t \langle \mathcal{K} \rangle \approx \partial_t \langle K^E \rangle$, and the connection between the Eulerian-mean and Lagrangian-mean viewpoints is straightforward: the kinetic energy extracted from the Eulerian-mean flow via Reynolds shear production approximately matches the creation of wave potential energy through buoyancy fluxes. In general, however, the connection is convoluted, involving the rate of change of the “Stokes energy” in (5.117).

4.7 Discussion and conclusions

4.7.1 Absence of a direct cascade of wave energy

The solutions reported here introduce the waves at $t = 0$ in (4.14) with infinite spatial scale. Wave refraction, $i\phi \zeta / 2$ in (4.8), immediately transfers wave energy to the smaller scales of the balanced relative vorticity ζ . This giant leap across wavenumbers is not a direct cascade of wave energy in the sense of Kolmogorov. And because of wave escape, the wave energy that is so efficiently transferred to eddy scales by refraction does not undergo a turbulence-driven direct cascade to the small length scales at which the wave dissipation in (4.9) is effective. This conclusion hinges on the assumption of a barotropic balanced flow and fixed vertical wavenumber. Similar to the vertical-plane-wave model, shallow-water models lack wave capture and an efficient direct cascade of wave energy [e.g., McIntyre 2009]. In less idealized models with balanced baroclinic shear, the vertical wavenumber can increase, and a direct cascade, perhaps resulting in wave capture, can ensue; stimulated generation might then be much stronger than in the vertical-plane-wave model considered here.

4.7.2 Regimes of wave-modified turbulence

Geostrophic straining accounts for most of the stimulated generation of wave energy in the examples considered in this paper. But refraction plays a fundamental role in these solutions with the uniform initial wave velocity in (4.14), because refraction creates the initial gradients of wave velocity that are then enhanced by geostrophic straining. We experimented by changing the initial condition of ϕ to an eddy-scale plane wave and repeated all the 2D turbulence solutions; the different initial condition significantly suppresses the initial refraction stage, but otherwise yields long-term solutions that are qualitatively similar to the solutions discussed above. Thus, to the extent that the uniform-wave initial condition (4.14) idealizes the generation of large-scale upper-ocean inertial oscillations by storms [e.g., Moehlis and Llewellyn Smith 2001; Danioux, Vanneste, and Bühler 2015], the initial refraction is a loss of lateral coherence, or a type of inertial pumping [Young and Ben-Jelloul 1997; Klein, Llewellyn Smith, and Lapeyre 2004], which is accompanied by an extraction of energy from the balanced

flow by the waves.

Although 10-20% of the balanced kinetic energy is converted into wave potential energy, and despite the wave breakage of the symmetry between cyclones and anti-cyclones, the wave-modified turbulence in section 4.4 remarkably resembles waveless two-dimensional turbulence [e.g., McWilliams 1984]: we still observe robust vortices and an increase in vortex length scale due to merger of like-signed vortices. Figure 4.7 shows small changes in the potential vorticity q and much larger changes in the wave PV q^w induced by changing the dispersivity. In this sense, the turbulent evolution is insensitive to wave modification.

To see significant wave modification of the turbulence we increased the amplitude of the initial wave in (4.14) so that $U_w = 6U_e$ (in section 4.4, $U_w = 2U_e$). With this level of wave energy the wave-modified 2D turbulence differs qualitatively from the waveless variety (not shown). The potential vorticity develops highly filamentary structures with little vortex formation; this inhibition of vortex formation is stronger in the weakly dispersive limit.

4.7.3 Energy transfers can be bi-directional in non-turbulent balanced flows

In all solutions considered in sections 4.2 and 4.4, the energy transfer is always from the balanced flow into the waves. This positive energy conversion, $\Gamma_r + \Gamma_a > 0$, remains true for very long turbulence simulations (not shown), because the refractive conversion is small at large time and turbulent stirring always increases lateral wave gradients: $\Gamma_r + \Gamma_a \approx \Gamma_a > 0$.

But the energy transfers can be bi-directional for non-turbulent flows. To illustrate this process, we consider a solution with initially uniform wave and $\psi(x, 0) = \sin x + \sin y$. This non-turbulent balanced flow consists of two vortices of opposite signs; there can be no inverse cascade from this initial condition because $\psi(x, 0)$ is already at the domain scale. Starting from the initial uniform wave, the refractive conversion generates wave potential energy \mathcal{P} at the expense of balanced kinetic energy \mathcal{K} (figure 4.14). After this initial refractive stage, however, \mathcal{P} and \mathcal{K} oscillate about equilibrium levels and there are quasi-periodic energy transfers back and forth between \mathcal{P} and \mathcal{K} .

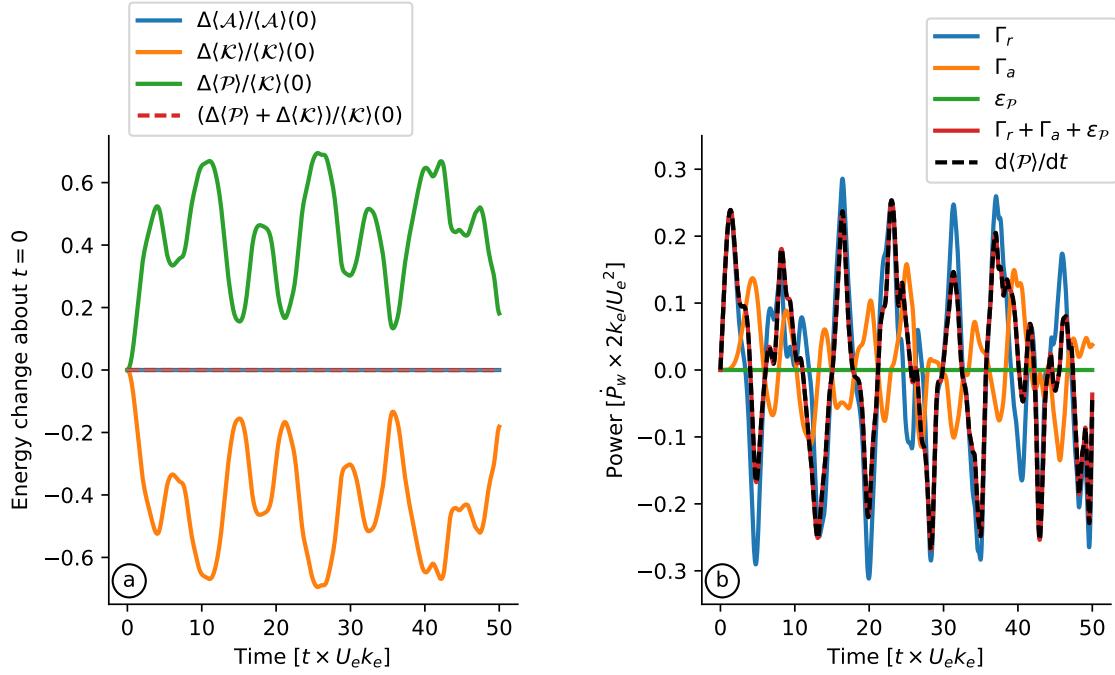


Figure 4.14: Diagnostics of an illustrative solution with initially uniform wave and $\psi(t = 0) = \sin x + \sin y$, with $\alpha = 0.1$ and $\hbar = 2$. (a) Energy difference about initial condition. (b) Wave potential energy budget.

In this solution, the refractive conversion Γ_r accounts for most of the quasi-periodic exchanges between \mathcal{K} and \mathcal{P} .

4.7.4 The correlation of wave amplitude with relative vorticity

A secondary result is the strong time-dependence of the correlation r , defined in (4.38), between incoherent waves and the relative vorticity; r measures the concentration of waves into cyclones or anti-cyclones [Danioux, Vanneste, and Bühler 2015]. Refraction concentrates waves into anti-cyclones and expels them from cyclones, thereby generating an initial strong negative r (see figure 4.8c). And Ehrenfest's theorem provides perhaps the simplest explanation for this concentration—see (4.31). As conjectured by Danioux, Vanneste, and Bühler [2015], the subsequent return of r towards zero (and even to positive values in the case $\hbar = 0.5$) is partially due to the unsteady geostrophic advection. The NIW-QG coupling compounds the unsteady advection: the dramatic initial concentration of waves into anti-cyclones weakens those vortices, with

ensuing development of positive skewness of relative vorticity (see figure 4.8d); the vorticity skewness increases with decreasing dispersivity because weakly dispersive waves extract more energy from the balanced flow (see figure 4.8).

4.7.5 Reconciling action conservation with RSP of wave kinetic energy

A main difficulty in making a connection between the NIW-QG model and earlier studies, such as those in table 1, is that the theory uses the Lagrangian-mean geostrophic flow as a primary variable. Previous studies generally employ an Eulerian wave-mean decomposition. Taylor and Straub [2016], for example, use frequency filtering to separate low frequency motions (Eulerian-mean) from high-frequency flow (dominated by near-inertial waves). Taylor & Straub find energy exchange between low-frequency (eddy) kinetic energy and high frequency (near-inertial) kinetic energy resulting from both vertical and horizontal Reynolds stresses; Taylor & Straub refer to this energy transfer as the “advection sink”, meaning a sink of low-frequency, Eulerian-mean kinetic energy. The Reynolds shear production RSP in (4.63) is analogous to the horizontal part of the advective sink.

The Reynolds stresses diagnostics of prior studies in table 4.1 might seem incompatible with fundamental aspects of the NIW-QG system. For example, the wave action \mathcal{A} in (4.15) differs from the wave kinetic energy $|\phi|^2/2$ only by the constant factor f_0 . Thus NIW-QG action conservation seems inconsistent with Reynolds stress transfer of kinetic energy from eddies to waves. But this impression is incorrect: Reynolds stresses aren’t inconsistent with conservation of \mathcal{A} , nor with stimulated generation.

To understand this, note that conservation of \mathcal{A} in (4.24) is a statement about the *leading order* wave velocities encoded in ϕ ; the balanced kinetic energy \mathcal{K} also involves *higher order* wave kinetic energy in the form of the “Stokes kinetic energy” \mathcal{K}^S : see (4.54) and the surrounding discussion. A main point from section 4.6.2 is that the energy transferred from the Eulerian-mean velocity by RSP is accounted for by \mathcal{K}^S (rather than the leading-order wave kinetic energy $f_0\mathcal{A}$).

At first glance it might seem that because stimulated generation results in an increase in wave potential energy \mathcal{P} , it must necessarily involve a transfer from eddy

potential energy. The vertical-plane-wave model is a counterexample: the eddies are barotropic and there is therefore no eddy potential energy to transfer. Indeed, if the eddies are barotropic, then *RSP* is the *only* pathway between the Eulerian-mean kinetic energy \mathcal{K}^E and wave energy. This rather obvious fact is deeply hidden by the Lagrangian-mean average: the *RSP* transfer out of \mathcal{K}^E passes first through \mathcal{K}^S on the way to \mathcal{P} . The intermediate passage through \mathcal{K}^S is an inevitable aspect of the Lagrangian-mean formulation of the NIW-QG system. All of these transfers can be diagnosed using Γ_a , Γ_r and Γ_s and (4.64) relates Reynolds shear production to the three Γ 's.

A full investigation of NIW-QG energetics, including diagnosis of Γ_r , Γ_a and Γ_s in solutions of the Boussinesq equations, is beyond the scope of this article. But we emphasize a crucial simplifying feature of the NIW-QG approximation: there are no important Stokes corrections to pressure and buoyancy, and therefore, to leading order, the Lagrangian-mean velocity is equal to the balanced velocity based on an Eulerian-mean pressure field. In principle, this is a straightforward way of diagnosing Lagrangian-mean velocities from numerical solutions of the Boussinesq equations.

4.7.6 Final remarks

There are many caveats to the application of our results to the post-storm oceanographic problem. Notably, the lack of geostrophic vertical shear suppresses important mechanisms of vertical refraction and straining, which introduce interesting modifications of the near-inertial wave physics [e.g., Thomas 2012] and can produce strong energy extraction by near-inertial waves [Shakespeare and Hogg 2017]. And our focus on quasi-inviscid initial value problems downplays the role of dissipation, but in forced-dissipative solutions, wave dissipation likely controls the strength of stimulated generation. Finally, better understanding the connection between numerical modeling studies [e.g. Taylor and Straub 2016] and the NIW-QG theory [XV; Wagner and Young 2016] deserves further investigation. We hope to explore these topics in future work.

Appendix 4.A: Details of the NIW-QG model

The NIW-QG model

Using multiscale asymptotic theory, Wagner and Young [2016] derive a model for the coupled evolution of QG balanced flow, near-inertial waves and their second harmonic. Assuming that the second harmonic is zero ($B = 0$), the Wagner and Young [2016] coupled model recovers the XV model in the limit where the waves have vertical scales much smaller than the balanced flow. In Wagner and Young [2016], the PV is

$$q = (\Delta + L)\psi + \beta y + \frac{1}{2f_0} [\Delta \frac{1}{2}|LA|^2 + iJ(LA^*, LA)], \quad (4.70)$$

where $\Delta \stackrel{\text{def}}{=} \partial_x^2 + \partial_y^2$ and $L \stackrel{\text{def}}{=} \partial_z(f_0/N)^2 \partial_z$, and LA is the back-rotated near-inertial velocity; the leading-order wave plus balanced flow velocity is

$$u + iv = LAe^{-if_0t} - \psi_y + i\psi_x, \quad (4.71)$$

where ψ is the streamfunction of the Lagrangian-mean (geostrophically balanced) flow. The PV is materially conserved,

$$q_t + J(\psi, q) = 0, \quad (4.72)$$

and the wave back-rotated velocity satisfies the YBJ equation,

$$LA_t + \frac{i}{2}f_0\Delta A + J(\psi, LA) + iLA(\frac{1}{2}\Delta\psi + \beta y) = 0. \quad (4.73)$$

The special family of solutions with a barotropic balanced flow $\psi = \psi(x, y, t)$, f -plane ($\beta = 0$), uniform background buoyancy frequency $N = N_0$, and $LA = e^{imz} \phi(x, y)$ wave velocity yields the vertical-plane-wave model in (4.7)-(4.8). The plane wave model is a solution of both XV and Wagner and Young [2016] equations because the barotropic flow assumption yields an infinite vertical-scale separation between waves and balanced flow.

Stokes drift

The horizontal component the Stokes drift of near-inertial waves is [Wagner and Young 2016]:

$$u^S + i\nu^S = \frac{i}{f_0} (M_{zz} \partial^* M^* - M_z^* \partial^* M_z), \quad (4.74)$$

$$w^S = \frac{i}{f_0} (M_z^* \partial \partial^* M - \partial M_z \partial^* M) + \text{cc}. \quad (4.75)$$

where $M = (f_0/N)^2 A_z$ and $M_z = LA$, and we recall $\partial = (\partial_x - i\partial_y)/2$. In the vertical-plane-wave model, the back-rotated velocity is $M_z = \phi e^{imz}$, and the horizontal Stokes drift above simplifies to

$$(u^S, \nu^S) = (\mathcal{A}_y, -\mathcal{A}_x), \quad (4.76)$$

where $\mathcal{A} = |\phi|^2/2f_0$ is the action density. Thus the Stokes drift of the vertical-plane-wave model is horizontally non-divergent, with streamfunction $-\mathcal{A}$. The vertical component of the Stokes drift reduces to

$$\begin{aligned} w^S &= \frac{1}{f_0 m} [\partial(\phi \partial^* \phi^*) + \partial^*(\phi^* \partial \phi)] \\ &= \frac{1}{m} [\Delta \frac{1}{2} \mathcal{A} - \eta^{-1} \nabla \cdot (\hat{k} \times \mathcal{F})], \end{aligned} \quad (4.77)$$

where we used (4.22) and the identities

$$\Delta = 4\partial \partial^* \quad \text{and} \quad J(f, g) = 2i(\partial^* f \partial g - \partial f \partial^* g). \quad (4.78)$$

An important property of the NIW-QG model is that there is no Lagrangian-mean vertical velocity:

$$w^E + w^S = 0. \quad (4.79)$$

Hence, using (5.54) and $\psi = \psi^E - \mathcal{A}$, we obtain an alternative expression for the potential vorticity (4.6) in terms of Eulerian-mean fields:

$$q = \Delta \psi^E + m w^E. \quad (4.80)$$

Appendix 4.B: Quadratic conservation laws

Ehrenfest's theorem

To obtain (4.25) we begin by noting that with \mathcal{F} defined in (4.20)

$$\partial_t \mathcal{F} = \frac{i}{2} \lambda^2 (\phi_t \nabla \phi^* - \phi^* \nabla \phi) + \frac{i}{4} \lambda^2 \nabla (\phi \phi_t^* - \phi^* \phi_t). \quad (4.81)$$

Multiplying the wave equation by $i\nabla\phi^*$, adding to the complex conjugate, and using identities such as

$$\nabla\phi^*J(\psi,\phi) - \nabla\phi J(\psi,\phi^*) = J(\phi^*,\phi)\nabla\psi, \quad (4.82)$$

one eventually finds

$$\begin{aligned} \partial_t \mathcal{F} - \frac{i}{4}\lambda^2 \nabla(\phi\phi_t^* - \phi^*\phi_t) + \frac{1}{4}\lambda^2\eta(\Delta\phi\nabla\phi^* + \Delta\phi^*\nabla\phi) = \\ - \nabla\psi \nabla \cdot (\hat{\mathbf{k}} \times \mathcal{F}) + \eta \frac{1}{2}\zeta \nabla \mathcal{A} + \frac{i}{2}\lambda^2(D_\phi \nabla\phi^* - D_{\phi^*} \nabla\phi). \end{aligned} \quad (4.83)$$

Taking the domain average, noting that the second and third terms on the left have zero average, and using the identity

$$\langle \nabla\psi \nabla \cdot (\hat{\mathbf{k}} \times \mathcal{F}) \rangle = \langle (\mathcal{F} \cdot \nabla) \hat{\mathbf{k}} \times \nabla\psi \rangle - \langle \hat{\mathbf{k}} \times (\zeta \mathcal{F}) \rangle, \quad (4.84)$$

we obtain (4.25) with the dissipative term

$$\varepsilon_{\mathcal{F}} \stackrel{\text{def}}{=} \frac{i}{2}\lambda^2 \langle D_\phi \nabla\phi^* - D_{\phi^*} \nabla\phi \rangle. \quad (4.85)$$

Wave potential energy

To obtain the wave potential energy equation (4.26) we take the dot product of $\nabla\phi^*$ with gradient of the wave equation (4.8) and add the complex conjugate; the calculation is best done using index notation. The final result is

$$\begin{aligned} \mathcal{P}_t + \nabla \cdot \left[\mathbf{u}^g \mathcal{P} + \frac{1}{2}\zeta \mathcal{F} + \frac{\lambda^2}{4} \frac{i}{2} \eta ((\nabla\phi \cdot \nabla)\nabla\phi^* - (\nabla\phi^* \cdot \nabla)\nabla\phi) \right] \\ = \frac{1}{2}\zeta \nabla \cdot \mathcal{F} - \frac{\lambda^2}{2} \phi_{,k} \sigma_{kl} \phi_{,l}^* + \frac{\lambda^2}{4} (\nabla\phi \cdot \nabla \mathcal{D}_{\phi^*} + \nabla\phi^* \cdot \nabla \mathcal{D}_{\phi}), \end{aligned} \quad (4.86)$$

where $\mathbf{u}^g \stackrel{\text{def}}{=} \hat{\mathbf{k}} \times \nabla\psi$ is the geostrophic velocity and

$$\sigma_{kl} \stackrel{\text{def}}{=} \frac{1}{2}(u_{k,l}^g + u_{l,k}^g) \quad (4.87)$$

is the geostrophic strain tensor. The local equation (4.86) integrates to (4.26), with the dissipative term

$$\varepsilon_{\mathcal{P}} = \frac{\lambda^2}{4} \langle \nabla\phi \cdot \nabla \mathcal{D}_{\phi^*} + \nabla\phi^* \cdot \nabla \mathcal{D}_{\phi} \rangle = -\frac{\lambda^2}{4} \langle \Delta\phi^* D_\phi + \Delta\phi \mathcal{D}_{\phi^*} \rangle. \quad (4.88)$$

Balanced kinetic energy

To obtain the balanced kinetic energy equation we first multiply the PV equation (4.7) by $-\psi$:

$$\mathcal{K}_t + \nabla \cdot [-\psi (\nabla \psi_t + \mathbf{u}^g q)] = \psi q_t^w - \psi \mathcal{D}_q. \quad (4.89)$$

To attack ψq_t^w on the right, we use the expression for q^w in (4.23). Thus

$$\psi q_t^w = \nabla \cdot \underbrace{\frac{1}{2} \left[\psi \left(\nabla \mathcal{A}_t + \frac{2}{\eta} \hat{\mathbf{k}} \times \mathcal{F}_t \right) - \nabla \psi \mathcal{A}_t \right]}_{\stackrel{\text{def}}{=} -\mathcal{H}_1} + \frac{1}{2} \zeta \mathcal{A}_t + \eta^{-1} \mathbf{u}^g \cdot \mathcal{F}_t. \quad (4.90)$$

Taking the dot product of (4.83) with \mathbf{u}^g we have

$$\begin{aligned} \mathbf{u}^g \cdot \mathcal{F}_t &= \nabla \cdot \underbrace{\left\{ \mathbf{u}^g \left[\frac{1}{2} \lambda^2 (\phi \phi_t^* - \phi^* \phi_t) + \frac{1}{4} \eta \lambda^2 |\nabla \phi|^2 \right] - \frac{1}{4} \eta \lambda^2 [\nabla \phi \mathbf{u}^g \cdot \nabla \phi^* + \nabla \phi^* \mathbf{u}^g \cdot \nabla \phi] \right\}}_{\stackrel{\text{def}}{=} -\eta \mathcal{H}_2} \\ &\quad + \eta \frac{1}{2} \zeta J(\psi, \mathcal{A}) + \frac{1}{2} \eta \lambda^2 \phi_{,k} \sigma_{kl} \phi_{,l}^* + \mathbf{u}^g \cdot \frac{i}{2} \lambda^2 (D_\phi \nabla \phi^* - \mathcal{D}_{\phi^*} \nabla \phi). \end{aligned} \quad (4.91)$$

Thus

$$\partial_t \mathcal{K} + \nabla \cdot [-\psi (\nabla \psi_t + \mathbf{u}^g q) + \mathcal{H}_1 + \mathcal{H}_2] = -\frac{1}{2} \zeta \nabla \cdot \mathcal{F} + \frac{1}{2} \lambda^2 \phi_{,k} \sigma_{kl} \phi_{,l}^* + \xi - \psi \mathcal{D}_q, \quad (4.92)$$

where

$$\xi = \frac{1}{2} f_0^{-1} (\phi^* D_\phi + \phi \mathcal{D}_{\phi^*}) \frac{1}{2} \zeta + f_0^{-1} \mathbf{u}^g \cdot \frac{i}{2} (\mathcal{D}_\phi \nabla \phi^* - \mathcal{D}_{\phi^*} \nabla \phi) \quad (4.93)$$

is the contribution of wave dissipation to the local balanced kinetic energy budget. Interestingly, the first term on the right of (4.93) reveals that the dissipation of wave action in anti-cyclones is a source of balanced kinetic energy. The second term on the right of (4.93) shows that the alignment of the “action-flux dissipation vector” $i(\mathcal{D}_\phi \nabla \phi^* - \mathcal{D}_{\phi^*} \nabla \phi)$, with the geostrophic velocity is also a source of balanced kinetic energy. The local equation (5.114) integrates to the balanced kinetic energy equation (4.27) with the dissipative terms

$$\Xi \stackrel{\text{def}}{=} \langle \xi \rangle \quad \text{and} \quad \varepsilon_{\mathcal{K}} = -\langle \psi \mathcal{D}_q \rangle. \quad (4.94)$$

Specific expressions with biharmonic dissipation

The dissipative terms in (4.7) and (4.8) add small dissipation to the energy equations in section 4.3. The wave kinetic energy dissipation added to (4.15) is

$$\varepsilon_{\mathcal{K}} = -\nu_w \langle |\Delta \phi|^2 \rangle. \quad (4.95)$$

The dissipation of wave potential energy in (4.26) is

$$\varepsilon_{\mathcal{P}} = -\frac{1}{2}\lambda^2 \nu_w \langle |\nabla \Delta \phi|^2 \rangle. \quad (4.96)$$

Similarly, the balanced kinetic energy dissipation in (4.27) is

$$\varepsilon_{\mathcal{K}} = \kappa_e \langle \psi \Delta^2 q \rangle = \kappa_e \langle q \Delta^2 \psi \rangle. \quad (4.97)$$

The wave dissipation contribution to the balanced kinetic energy budget is

$$\Xi = \frac{1}{2} \nu_w f_0^{-1} \left\langle \frac{1}{2} \zeta (\phi^* \Delta^2 \phi + \phi \Delta^2 \phi^*) \right\rangle + \nu_w f_0^{-1} \left\langle \frac{i}{2} \psi [J(\phi^*, \Delta^2 \phi) - J(\phi, \Delta^2 \phi^*)] \right\rangle. \quad (4.98)$$

In all solutions of initial value problems reported in this paper, the dissipative terms (4.95), (4.96), (4.97), and (4.93) account for less—typically much less—than 10% of the energy tendencies.

We thank three anonymous referees for constructive criticism and suggestions. This study was supported by the National Aeronautics and Space Administration (NNX-16AO5OH) and the National Science Foundation (OCE-1357047 and OCE-1657041).

Chapter 4, in full, reprints material as it appears in Journal of Fluid Mechanics, in press, doi:10.1017/jfm.2018.308. Rocha, C. B.; Wagner, G. L.; Young, W. R. The dissertation author was the primary investigator and author of this paper.

Chapter 5

Equilibration of forced 2D turbulence by stimulated generation

5.1 Introduction

Motivated by the work of Gertz and Straub [2009], who studied the equilibration of a forced 2D doubled-gyre flow via energy transfers to 3D near-inertial waves, and by an increasing body of forced-dissipative Boussinesq simulations driven by both low-frequency and high-frequency forcing [e.g., Taylor and Straub 2016; Barkan, Winters, and McWilliams 2016], this chapter aims to investigate the equilibration of forced 2D turbulence through stimulated generation in the vertical plane wave model. Because the effects of dissipation on the potential vorticity are non-obvious, the first step in this program is to re-derive the vertical plane wave NIW-QG model of Xie and Vanneste [2015] (hereafter XV15) and Wagner and Young [2016] (hereafter WY16) starting from the dissipative Boussinesq equations.

Derived in detail in section 5.3, the dissipative NIW-QG model describes the coupled evolution of near-inertial waves with velocity

$$\tilde{u} + i\tilde{v} = \phi(x, y, t)e^{i(mz - f_0 t)}, \quad (5.1)$$

and a Lagrangian-mean barotropic quasi-geostrophic flow with streamfunction $\psi^L = \bar{p}/f_0$, where \bar{p} is both an Eulerian-mean and a Lagrangian-mean pressure [Rocha, Wagner, and Young 2018]. Waves and quasi-geostrophic flow are coupled through the

potential vorticity (PV):

$$q = \underbrace{\Delta \psi^L}_{\stackrel{\text{def}}{=} \zeta} + \frac{1}{2} \Delta \mathcal{A} + \eta^{-1} \nabla \cdot (\hat{\mathbf{z}} \times \mathcal{F}), \quad (5.2)$$

where \mathcal{A} is the wave action density,

$$\mathcal{A} \stackrel{\text{def}}{=} \frac{1}{2f_0} |\phi|^2, \quad (5.3)$$

and \mathcal{F} is the wave action flux,

$$\mathcal{F} \stackrel{\text{def}}{=} \frac{i}{4} \lambda^2 (\phi \nabla \phi^* - \phi^* \nabla \phi). \quad (5.4)$$

Also in (5.2), the horizontal Laplacian is

$$\Delta \stackrel{\text{def}}{=} \partial_x^2 + \partial_y^2. \quad (5.5)$$

The streamfunction ψ^L in (5.2) is

$$\psi^L = \psi^E - \mathcal{A}, \quad (5.6)$$

where $-\mathcal{A}$ is the streamfunction of the Stokes drift and ψ^E is the streamfunction of the Eulerian-mean flow.

The wave amplitude ϕ satisfies the dissipative Young and Ben-Jelloul [1997] (hereafter YBJ) equation:

$$\phi_t + J(\psi^L, \phi) + \frac{i}{2} \zeta \phi - \frac{i}{2} \eta \Delta \phi = -\gamma \phi, \quad (5.7)$$

where $\eta \stackrel{\text{def}}{=} N^2/f_0 m^2$ is the “wave dispersivity,” and $\gamma = \mu + m^2 \nu$, with μ the Rayleigh friction coefficient, ν the viscosity, and m the vertical wavenumber in (5.1). Isotropic viscosity in the parent Boussinesq equations implies linear dissipation in the YBJ equation (5.7). Also in (5.7), the Jacobian is

$$J(f, g) \stackrel{\text{def}}{=} f_x g_y - f_y g_x. \quad (5.8)$$

The potential vorticity (5.2) satisfies

$$q_t + J(\psi^L, q) = D_q, \quad (5.9)$$

where

$$D_q \stackrel{\text{def}}{=} -\mu\zeta + (\gamma - \mu)\Delta\mathcal{A} - 2\frac{\gamma}{\eta}\nabla \cdot (\hat{\mathbf{z}} \times \mathcal{F}); \quad (5.10)$$

see (5.89) for equivalent expression of D_q . In contrast to the dissipation of wave back-rotated velocity on the right of (5.7), D_q is non-monotonic, so dissipation can generate PV locally [c.f., Herring, Kerr, and Rotunno 1994]; the horizontally averaged PV $\langle q \rangle$ is conserved because D_q integrates to zero $\langle D_q \rangle = 0$. Section 5.4 illustrates this process using an initial-value problem where PV is generated locally by wave dissipation, despite being initially zero everywhere.

The non-obvious dissipative term D_q in (5.9)-(5.10) ensures that dissipation has a negative-definite effect on the Eulerian-mean kinetic energy,

$$\partial_t \left\langle f_0 \mathcal{A} + \underbrace{\frac{1}{2} |\nabla \psi^E|^2}_{\stackrel{\text{def}}{=} \mathcal{K}^E} \right\rangle = \text{RSP} - \mu \langle |\nabla \psi^E|^2 \rangle - 2\gamma f_0 \langle \mathcal{A} \rangle, \quad (5.11)$$

where RSP is the Reynolds shear production. Thus the asymptotic model (5.2)-(5.10) preserves a key property of the dissipative Boussinesq equations. The “Lagrangian-mean kinetic energy,”

$$\mathcal{K}^L \stackrel{\text{def}}{=} \frac{1}{2} |\nabla \psi^L|^2, \quad (5.12)$$

does not decay monotonically due to dissipation, and there is no reason to expect that it should. In fact, wave dissipation can generate \mathcal{K}^L , and this “wave streaming” is a non-negligible source of Lagrangian-mean kinetic energy in all solutions considered in this chapter.

Using the dissipative NIW-QG model above, we obtain forced solutions of the vertical plane wave model in order to study the equilibration of 2D turbulence by stimulated generation of near-inertial waves. The turbulence is driven by isotropic stochastic forcing at wavenumber $k_f L / 2\pi = 8$; waves with infinite horizontal scales are also forced stochastically. Thanks to stimulated generation, forced 2D turbulence equilibrates even without bottom drag. And as in the problem of freely decaying 2D turbulence modified by near-inertial waves [Rocha, Wagner, and Young 2018], geostrophic straining accounts for most of stimulated generation.

5.2 The dissipative Boussinesq equations

We start from the hydrostatic Boussinesq equations with Rayleigh friction and isotropic viscosity and diffusivity [e.g., Vallis 2006]:

$$u_t + \mathbf{u} \cdot \nabla u - f_0 v = -p_x - \mu u + \nu \nabla^2 u, \quad (5.13a)$$

$$v_t + \mathbf{u} \cdot \nabla v + f_0 u = -p_y - \mu v + \nu \nabla^2 v, \quad (5.13b)$$

$$0 = -p_z + b, \quad (5.13c)$$

$$b_t + \mathbf{u} \cdot \nabla b + w N_0^2 = \kappa \nabla^2 b, \quad (5.13d)$$

$$\nabla \cdot \mathbf{u} = 0, \quad (5.13e)$$

where the velocity is $\mathbf{u} = u \hat{\mathbf{x}} + v \hat{\mathbf{y}} + w \hat{\mathbf{z}}$, and the linear operators are

$$\nabla \stackrel{\text{def}}{=} \partial_x \hat{\mathbf{x}} + \partial_y \hat{\mathbf{y}} + \partial_z \hat{\mathbf{z}}, \quad \text{and} \quad \nabla^2 \stackrel{\text{def}}{=} \nabla \cdot \nabla = \partial_x^2 + \partial_y^2 + \partial_z^2. \quad (5.14)$$

Rayleigh friction with coefficient μ in (5.13a) and (5.13b) yields a linear-drag term in the potential vorticity equation, which parameterizes boundary layer turbulence that damps large-scale barotropic flow.

The hydrostatic approximation is valid provided the aspect ratio of the flow is small,

$$\alpha \stackrel{\text{def}}{=} \frac{H}{L} \ll 1, \quad (5.15)$$

a condition easily satisfied by geostrophic flow and near-inertial waves. For convenience, we make the f-plane approximation, and we assume a uniform background stratification so that the density is

$$\rho(\mathbf{x}, t) = \rho_0 [1 - g^{-1} N_0^2 z - g^{-1} b(\mathbf{x}, t)], \quad (5.16)$$

ρ_0 is a reference density, N^2 is the stratification frequency, and b is the dynamical buoyancy perturbation.

Without dissipation, the potential vorticity (PV)

$$\Pi \stackrel{\text{def}}{=} (f_0 \hat{\mathbf{z}} + \boldsymbol{\omega}) \cdot (N_0^2 \hat{\mathbf{z}} + \nabla b), \quad (5.17)$$

with the “hydrostatic vorticity”

$$\boldsymbol{\omega} = -v_z \hat{\mathbf{x}} + u_z \hat{\mathbf{y}} + \underbrace{(v_x - u_y)}_{\stackrel{\text{def}}{=} \omega} \hat{\mathbf{z}}, \quad (5.18)$$

is conserved on fluid particles [e.g., Pedlosky 1987]. To form an equation for Π we dot $N^2\hat{\mathbf{z}} + \nabla b$ with the curl of (5.13a)-(5.13c), and add the result to the dot product between $f_0\hat{\mathbf{z}} + \boldsymbol{\omega}$ and the gradient of (5.13d). The resulting dissipative PV equation is

$$\Pi_t + \mathbf{u} \cdot \nabla \Pi = D_\Pi, \quad (5.19)$$

where

$$D_\Pi \stackrel{\text{def}}{=} -\mu(\nabla b + N_0^2\hat{\mathbf{z}}) \cdot \boldsymbol{\omega} + \nu(\nabla^2 \boldsymbol{\omega}) \cdot (\nabla b + N_0^2\hat{\mathbf{z}}) + \kappa(\nabla^2 \nabla b) \cdot (f_0\hat{\mathbf{z}} + \boldsymbol{\omega}) \quad (5.20)$$

[c.f., Pedlosky 1987]. The dissipative terms in D_Π are non-monotonic, so PV can be generated by dissipation [e.g., Herring, Kerr, and Rotunno 1994]. Like D_Π , the reduced D_q derived below is non-monotonic, with the important consequence that wave dissipation can generate balanced PV q . Section 5.4 discusses a dramatic example where wave dissipation generates q from an initial condition with $q = 0$ everywhere.

5.2.1 The non-dimensional two-time equations

To obtain the dissipative QG-NIW vertical plane wave model, we follow WY16's two-time derivation, expanding both primitive equations (5.13a)-(5.13e) and PV equation (5.17)-(5.19) in powers of wave non-linearity ϵ and phase-averaging over fast time. The algebra is dramatically simplified using YBJ's notation, with complex horizontal velocity \mathcal{U} ,

$$\mathcal{U} \stackrel{\text{def}}{=} u + i\nu, \quad (5.21)$$

and complex spatial coordinate s ,

$$s \stackrel{\text{def}}{=} x + iy, \quad \text{so that} \quad \partial_s = \frac{1}{2}(\partial_x - i\partial_y). \quad (5.22)$$

Two useful complex differential identities used throughout sections 5.2 and 5.3 are the horizontal Laplacian

$$\Delta \stackrel{\text{def}}{=} \partial_x^2 + \partial_y^2 = 4\partial_s \partial_{s^*}, \quad (5.23)$$

and the Jacobian

$$J(f, g) \stackrel{\text{def}}{=} f_x g_y - f_y g_x = 2i(f_{s^*} g_s - f_s g_{s^*}). \quad (5.24)$$

Primitive equations

Using the YBJ complex notation and WY16's scaling, the dissipative Boussinesq equations (5.13a)-(5.13e) non-dimensionalize to

$$\mathcal{U}_{\tilde{t}} + i f_0 \mathcal{U} = -\epsilon (\mathbf{u} \cdot \nabla \mathcal{U} + 2p_{s^*}) - \epsilon^2 [\mathcal{U}_{\tilde{t}} + \mu \mathcal{U} - \nu \mathcal{U}_{zz}] + \mathcal{O}(\alpha^2 \epsilon^2), \quad (5.25a)$$

$$p_z = \epsilon b, \quad (5.25b)$$

$$b_{\tilde{t}} + wN^2 = -\epsilon \mathbf{u} \cdot \nabla b - \epsilon^2 [b_{\tilde{t}} - \kappa b_{zz}] + \mathcal{O}(\alpha^2 \epsilon^2), \quad (5.25c)$$

$$\mathcal{U}_s + \mathcal{U}_{s^*}^* + w_z = 0. \quad (5.25d)$$

In (5.25) the coefficient ϵ indicates that ϵ -terms are $\mathcal{O}(\epsilon)$ compared to the leading order terms—we keep the physical parameters in (5.25) and throughout to improve readability. The small parameter ϵ is the wave non-linearity,

$$\epsilon \stackrel{\text{def}}{=} \frac{\tilde{U}}{f_0 L}, \quad (5.26)$$

and

$$\partial_t = \partial_{\tilde{t}} + \epsilon^2 \partial_{\tilde{t}}. \quad (5.27)$$

Implicit in the scaling above is that the balanced velocity is $\bar{U} = \epsilon \tilde{U}$, so that

$$\text{Ro} \stackrel{\text{def}}{=} \frac{\bar{U}}{f_0 L} = \epsilon^2. \quad (5.28)$$

Introduced by XV15 and WY16, this convenient “strong-wave” assumption promotes the wave effects to algebraically accessible orders in the expansion (see appendix 5.A). But the limit of $\tilde{U}/\bar{U} \ll 1$ and $\text{Ro} \ll 1$ is recovered by the ensuing reduced model; see Rocha, Wagner, and Young [2018] for further discussion.

Also in (5.25), we assume that the dissipative small parameters are

$$\frac{\mu}{f_0} = \epsilon^2 \quad \text{and} \quad \frac{\nu}{f_0 H^2} = \epsilon^2. \quad (5.29)$$

With this choice, dissipation appears at the solvability order, thus entering in the YBJ equation. For the diffusivity, we assume $\nu \sim \kappa$, so that

$$\frac{\kappa}{f_0 H^2} = \epsilon^2. \quad (5.30)$$

Because of the small aspect ratio $\alpha \ll 1$, only the ∂_z^2 part of the 3D Laplacian ∇^2 appears in (5.25).

Potential vorticity

The perturbation of PV about the background, rest PV $f_0 N_0^2$ is also conserved on fluid particles. Introducing this ‘available PV’ [Wagner and Young 2015] simplifies the asymptotics below. A convenient definition is

$$Q \stackrel{\text{def}}{=} \frac{\Pi}{N_0^2} - f_0 = \omega + \frac{b_z}{N_0^2} + \epsilon \frac{\boldsymbol{\omega} \cdot \nabla b}{N_0^2}, \quad (5.31)$$

where we recall $\boldsymbol{\omega} = \hat{\mathbf{z}} \cdot \boldsymbol{\omega} = v_x - u_y$. The non-dimensional PV satisfies

$$Q_{\tilde{t}} + \epsilon \boldsymbol{u} \cdot \nabla Q + \epsilon^2 Q_{\tilde{t}} = \epsilon^2 D_Q, \quad (5.32)$$

with

$$D_Q = -\mu \omega + \nu \omega_{zz} + \kappa \frac{f_0}{N^2} b_{zzz} - \frac{\epsilon}{N^2} (\mu \boldsymbol{\omega} - \nu \boldsymbol{\omega}_{zz}) \cdot \nabla b + \epsilon \frac{\kappa}{N^2} \boldsymbol{\omega} \cdot \nabla b_{zz} + \mathcal{O}(\alpha^2 \epsilon^2). \quad (5.33)$$

Again only Rayleigh damping and vertical small-scale viscous and diffusive dissipation appear at the lowest orders in (5.33).

5.3 The expansion

Following WY16’s program, we expand the dynamical variables in powers of ϵ ,

$$\mathcal{U} = \mathcal{U}_0 + \epsilon \mathcal{U}_1 + \epsilon^2 \mathcal{U}_2 + \dots, \quad (5.34)$$

$$\boldsymbol{\omega} = \boldsymbol{\omega}_0 + \epsilon \boldsymbol{\omega}_1 + \epsilon^2 \boldsymbol{\omega}_2 + \dots, \quad (5.35)$$

etc, introduce these expansions into (5.25)-(5.33), and collect same-order terms.

5.3.1 ϵ^0 : near-inertial waves

The ϵ^0 equations govern the dynamics of free hydrostatic near-inertial waves:

$$\mathcal{U}_{0\tilde{t}} + i f_0 \mathcal{U}_0 = 0, \quad (5.36a)$$

$$0 = p_{0z}, \quad (5.36b)$$

$$\mathcal{U}_{0s} + \mathcal{U}_{0s^*}^* + w_{0z} = 0, \quad (5.36c)$$

$$b_{0\tilde{t}} + w_0 N_0^2 = 0. \quad (5.36d)$$

While the leading-order equations (5.36) admit a general YBJ solution, we focus on a single plane wave in the vertical and barotropic Eulerian-mean flow [Xie and Vanneste 2015; Rocha, Wagner, and Young 2018]:

$$\mathcal{U}_0 = \phi e^{i\varpi}, \quad (5.37a)$$

$$w_0 = im^{-1}(\phi_s e^{i\varpi} - \phi_{s^*}^* e^{-i\varpi}), \quad (5.37b)$$

$$b_0 = m\eta(\phi_s e^{i\varpi} + \phi_{s^*}^* e^{-i\varpi}), \quad (5.37c)$$

$$p_0 = p_0(x, y, \bar{t}), \quad (5.37d)$$

where $\varpi = mz - f_0 t$ is the phase of the vertical plane wave. Appendix 5.B proves a useful identity of this leading-order solution:

$$m\mathbf{u}_0 = -\boldsymbol{\omega}_0, \quad (5.38)$$

from which follows several other identities used below. The leading-order equations (5.36) also imply that

$$N_0^2 \omega_0 = -f_0 b_{0z}, \quad (5.39)$$

so the leading-order PV vanishes $Q_0 = 0$. This is the standard statement that “internal waves have no available PV.” But internal waves do have non-linear PV, which enters at higher orders in the expansion [Wagner and Young 2015].

5.3.2 ϵ^1 : wave-averaged geostrophic flow

The ϵ^1 equations are

$$\mathcal{U}_{1\tilde{t}} + if_0 \mathcal{U}_1 = -\mathbf{u}_0 \cdot \nabla \mathcal{U}_0 - 2p_{0s^*}, \quad (5.40)$$

$$p_{1z} = b_0, \quad (5.41)$$

$$\mathcal{U}_{1s} + \mathcal{U}_{1s^*}^* + w_{1z} = 0, \quad (5.42)$$

$$b_{1\tilde{t}} + w_1 N_0^2 = -\mathbf{u}_0 \cdot \nabla b_0. \quad (5.43)$$

To work on (5.40) we note the identity

$$\begin{aligned} \mathbf{u}_0 \cdot \nabla \mathcal{U}_0 &= \mathcal{U}_0 \mathcal{U}_{0s} + \mathcal{U}_0^* \mathcal{U}_{0s^*} + w_0 \mathcal{U}_{0z} \\ &= \phi \phi_s e^{2i\varpi} + \phi^* \phi_{s^*} - (\phi_s e^{i\varpi} - \phi_{s^*}^* e^{-i\varpi}) \phi e^{i\varpi} \\ &= \phi^* \phi_{s^*} + \phi \phi_{s^*}^* = 2f_0 \mathcal{A}_{s^*}, \end{aligned} \quad (5.44)$$

where \mathcal{A} is the near-inertial action density:

$$\mathcal{A} \stackrel{\text{def}}{=} \frac{|\phi|^2}{2f_0}. \quad (5.45)$$

An interesting property of the vertical plane wave model is that the $e^{2i\omega}$ terms cancel out, and the non-linear term $\mathbf{u}_0 \cdot \nabla \mathcal{U}_0$ is written in terms of the action gradient \mathcal{A}_{s^*} .

Wave-modified geostrophic balance

The self-averaging property of $\mathbf{u}_0 \cdot \nabla \mathcal{U}_0$, together with the leading-order pressure's independence of fast time, imply that the solution of the first-order horizontal momentum equation (5.40) is

$$\mathcal{U}_1 + \mathcal{U}^S = 2i\psi_{s^*}^L, \quad (5.46)$$

where ψ^L is the geostrophic streamfunction,

$$\psi^L = f_0^{-1} p_0, \quad (5.47)$$

and \mathcal{U}^S is the Stokes drift,

$$\mathcal{U}^S = \overline{\xi_0 \cdot \nabla \mathcal{U}_0} = -if_0^{-1} \overline{\xi_0 \cdot \nabla \mathcal{U}_{0\tilde{t}}} \quad (5.48)$$

$$= if_0^{-1} \overline{\mathbf{u}_0 \cdot \nabla \mathcal{U}_{0\tilde{t}}} \quad (5.49)$$

$$= -2i\mathcal{A}_{s^*}, \quad (5.50)$$

where ξ_0 is the Lagrangian displacement, $\mathbf{u}_0 = \xi_{0\tilde{t}}$, and we have used the identity (5.44).

The streamfunction in the wave-geostrophic relation (5.46) is associated with the Lagrangian-mean—Eulerian-mean plus Stokes drift—flow [Wagner and Young 2015]. The pressure in (5.47) is both an Eulerian and a Lagrangian mean, because for near-inertial waves the Stokes pressure correction is insignificant. A simplifying property of the vertical plane wave model is that the Stokes drift is horizontally non-divergent [Rocha, Wagner, and Young 2018], so the Eulerian-mean horizontal velocity is

$$\mathcal{U}_1 = 2i(\underbrace{\psi^L + \mathcal{A}}_{\stackrel{\text{def}}{=} \psi^E})_{s^*}, \quad (5.51)$$

where ψ^E is the streamfunction of the Eulerian-mean flow.

Additional ϵ^1 fields

Using the leading-order buoyancy (5.37c) in the hydrostatic relation (5.41) yields the near-inertial wave pressure:

$$p_1 = -i\eta(\phi_s e^{i\omega} - \phi_{s*}^* e^{-i\omega}). \quad (5.52)$$

The vertical velocity w_1 is also independent of fast time, because $\mathcal{U}_1 = \overline{\mathcal{U}}_1$. Phase-averaging the buoyancy equation (5.43) implies that the Lagrangian-mean vertical velocity vanishes [Wagner and Young 2015]:

$$w_1 + w^S = 0, \quad (5.53)$$

where w^S is the Stokes vertical velocity,

$$\begin{aligned} N_0^2 w^S &= N_0^{-2} \overline{\xi_0 \cdot \nabla w_0} \\ &= -\overline{\xi_0 \cdot \nabla b_{0\tilde{t}}} = \overline{u_0 \cdot \nabla b_0}. \end{aligned} \quad (5.54)$$

Plugging the leading-order solution (5.37) in (5.53)-(5.54) yields a useful expression for w_1 in terms of the back-rotated wave velocity ϕ :

$$\begin{aligned} mw_1 &= -\frac{1}{f_0} (\phi \phi_{ss*}^* + \phi^* \phi_{ss*} + 2|\phi_s|^2) \\ &= -\frac{1}{2} \Delta \mathcal{A} + \frac{i}{2f_0} J(\phi^*, \phi), \end{aligned} \quad (5.55)$$

where the horizontal Laplacian and Jacobian operators are defined in (5.23) and (5.24). Because w_1 is independent of fast time, the buoyancy equation reduces to

$$b_{1\tilde{t}} = -(u_0 \cdot \nabla b_0 - \overline{u_0 \cdot \nabla b_0}) \quad (5.56)$$

$$= m\eta(\frac{1}{2}\phi^2)_{ss} e^{2i\omega} + CC, \quad (5.57)$$

with solution

$$b_1 = \frac{i}{2} m\eta f_0^{-1} (\frac{1}{2}\phi^2)_{ss} e^{2i\omega} + CC. \quad (5.58)$$

The ϵ^1 PV is

$$Q_1 = \omega_1 + N_0^{-2} \overline{\omega_0 \cdot \nabla b_0}, \quad (5.59)$$

where

$$\omega_0 = -v_{0z} \hat{x} + u_{0z} \hat{y} + (v_{0x} - u_{0y}) \hat{z}; \quad (5.60)$$

the ϵ^1 PV equation (5.32) implies that Q_1 is independent of fast time. Using the identity in (5.38) in (5.54) yields yet another useful identity:

$$\begin{aligned} \overline{\omega_0 \cdot \nabla b_0} &= -m N_0^2 w^S \\ &= m N_0^2 w_1. \end{aligned} \quad (5.61)$$

Thus Q_1 in (5.59) becomes

$$Q_1 = \Delta\psi^E + mw_1 \quad (5.62)$$

$$= \Delta\psi^L + \frac{1}{2}\Delta\mathcal{A} + \frac{i}{2f_0}J(\phi^*, \phi). \quad (5.63)$$

5.3.3 ϵ^2 : amplitude equation and second-order corrections

The ϵ^2 equations are

$$\mathcal{U}_{2\bar{t}} + i f_0 \mathcal{U}_2 = -\mathbf{u}_0 \cdot \nabla \mathcal{U}_1 - \mathbf{u}_1 \cdot \nabla \mathcal{U}_0 - 2p_{1s^*} - \mathcal{U}_{0\bar{t}} - (\mu - v\partial_z^2) \mathcal{U}_0, \quad (5.64a)$$

$$0 = p_{2z} + b_1, \quad (5.64b)$$

$$\mathcal{U}_{2s} + \mathcal{U}_{2s^*}^* + w_{2z} = 0, \quad (5.64c)$$

$$b_{2t} + w_2 N_0^2 = -\mathbf{u}_0 \cdot \nabla b_1 - \mathbf{u}_1 \cdot \nabla b_0 - b_{0\bar{t}} + \kappa b_{0zz}. \quad (5.64d)$$

These equations differ from those in YBJ and WY16 by dissipative terms on the right of (5.64a) and (5.64d), which we scaled to appear at this order.

Amplitude equation

On the right of (5.64a), the second-order horizontal velocity \mathcal{U}_2 is forced both by resonant terms, proportional to $e^{i\omega}$, and non-resonant terms, proportional to $e^{-i\omega}$. An equation for ϕ follows from removing the resonant terms from (5.64). Working on the first term on the right of (5.64), we find

$$u_0 \cdot \nabla \mathcal{U}_1 = \phi \mathcal{U}_{1s} e^{i\omega} + \phi^* \mathcal{U}_{1s^*} e^{-i\omega} \quad (5.65)$$

$$= \underbrace{(2i\psi_{ss^*}^L \phi + 2i\mathcal{A}_{ss^*} \phi)}_{= \frac{i}{2}\zeta\phi} e^{i\omega} + (2i\psi_{s^*s^*}^E \phi) e^{-i\omega}, \quad (5.66)$$

where we have used the identities of appendix 5.A. Next, the second term on the right of (5.64a) is

$$\mathbf{u}_1 \cdot \nabla \mathcal{U}_0 = \underbrace{[2i(\psi_{s^*}^L \phi_s - \psi_s^L \phi_{s^*})]}_{=J(\psi^L, \phi)} + \underbrace{2i(\mathcal{A}_{s^*} \phi_s - \mathcal{A}_s \phi_{s^*})}_{=J(\mathcal{A}, \phi)} + imw_1 \phi e^{i\omega}. \quad (5.67)$$

And the pressure term is

$$2p_{1s^*} = -\underbrace{2i\eta\phi_{ss^*}}_{=\frac{i}{2}\eta\Delta\phi} e^{i\omega} + 2i\eta\phi_{s^*s^*}^* e^{-i\omega}. \quad (5.68)$$

WY16 remark that the resonant cubic terms in (5.64) add to zero. This cancellation is easy to show in the vertical plane-wave model with the identity

$$-imw_1 \phi = \frac{i}{2}\phi \Delta \mathcal{A} + J(\mathcal{A}, \phi) \quad (5.69)$$

that follows from (5.55). Thus the solvability condition—the requirement that the resonant terms in (5.64) add to zero—yields the dissipative generalization of the YBJ amplitude equation:

$$\phi_t + J(\psi^L, \phi) + \frac{i}{2}\zeta\phi - \frac{i}{2}\eta\Delta\phi = -\underbrace{(\mu + m^2\nu)}_{\stackrel{\text{def}}{=} \gamma} \phi, \quad (5.70)$$

where we recall that $\eta = f_0\lambda^2$ and $\lambda = N_0/f_0m$. Wave dissipation on the right of (5.70) is linear because, in the vertical plane wave model, vertical viscosity yields $\nu\partial_z^2 \mapsto -m^2\nu$, so in this model Laplacian viscosity yields a linear dissipation of the back-rotated near-inertial velocity ϕ , which adds to the Rayleigh friction μ .

Second-order horizontal velocity

After removing the resonant terms, the second-order momentum equation (5.64a) becomes (c.f., YBJ)

$$\mathcal{U}_{2t} + if_0 \mathcal{U}_2 = -2i(\psi_{s^*s^*}^E \phi^* + \eta\phi_{s^*s^*}^*) e^{-i\omega}, \quad (5.71)$$

with solution

$$\mathcal{U}_2 = -\underbrace{[\lambda^2\phi_{s^*s^*}^* + f_0^{-1}\psi_{s^*s^*}^E \phi^*]}_{\stackrel{\text{def}}{=} \phi_2^*} e^{-i\omega}, \quad (5.72)$$

where we recall that $\psi^E = \psi^L + \mathcal{A}$ is the streamfunction of the Eulerian-mean flow defined in (5.51). The second-order correction induces a slight distortion of the wave velocity hodograph (the leading-order wave velocity is circularly polarized). This distortion is crucial for non-zero phase-averaged wave Reynolds stresses that lead to wave energy extraction from the Eulerian-mean flow [Thomas 2012; Rocha, Wagner, and Young 2018].

Second-order potential vorticity equation

Using properties of the leading-order solution, such as the Beltrami identity in (5.141), the second-order PV equation is

$$Q_{2\bar{t}} + \mathbf{u}_0 \cdot \nabla Q_1 = m(\gamma - m^2\kappa)w_0, \quad (5.73)$$

where we recall that $\gamma = \mu + m^2\nu$. Using leading-order buoyancy equation (5.37c), and noting that Q_1 is independent of fast time, the solution of (5.73) follows from fast-time integration:

$$Q_2 = -\boldsymbol{\xi}_0 \cdot \nabla Q_1 - \frac{m}{N_0^2}(\gamma - m^2\kappa)b_0 + \bar{Q}_2, \quad (5.74)$$

where \bar{Q}_2 is an inconsequential constant of integration [Wagner and Young 2015]. Note the appearance of a dissipative term, proportional to b_0 in (5.37c).

5.3.4 ϵ^3 : geostrophic PV advection

With the identity (5.61), the phase-averaged ϵ^3 PV equation is

$$Q_{1\bar{t}} + \mathbf{u}_1 \cdot \nabla Q_1 + \overline{\mathbf{u}_0 \cdot \nabla Q_2} = -\mu\omega_1 + (\gamma + m^2\kappa)mw_1. \quad (5.75)$$

Because Q_1 is independent of fast time, and using the intermediate result

$$\mathbf{u}^S = \overline{\boldsymbol{\xi}_0 \cdot \nabla \mathbf{u}_0} = -\overline{\mathbf{u}_0 \cdot \nabla \boldsymbol{\xi}_0} \quad (5.76)$$

with (5.74) and (5.61)-(5.38), yields

$$\overline{\mathbf{u}_0 \cdot \nabla Q_2} = \mathbf{u}^S \cdot \nabla Q_1 - (\gamma - m^2\kappa)mw_1. \quad (5.77)$$

Noting the cancellation of the κ -terms in (5.77) and (5.78), the PV equation (5.78) reduces to

$$Q_{1\bar{t}} + \mathbf{u}^L \cdot \nabla Q_1 = -\mu\omega_1 - 2\gamma mw_1, \quad (5.78)$$

where recall that

$$\omega_1 = v_{1x} - u_{1y} = \Delta\psi^E, \quad (5.79)$$

and w_1 is given by (5.55). The wave-modified PV, Q_1 in (5.59), is advected by the Lagrangian-mean flow [Wagner and Young 2015].

5.3.5 Summary of asymptotics

In summary, the asymptotic QG-NIW model consists of two evolutions equations, one for back-rotated wave velocity (5.70) and the other for the wave-averaged balanced PV (5.78). And the elliptic PV equation (5.59) couples waves and balanced flow. Using the notation of Rocha, Wagner, and Young [2018] and the identity

$$\nabla \cdot (\hat{\mathbf{z}} \times \mathcal{F}) = \frac{i}{2} \lambda^2 J(\phi^*, \phi), \quad (5.80)$$

the PV in (5.59) becomes

$$q = \underbrace{\Delta\psi^L}_{\stackrel{\text{def}}{=} \zeta} + \underbrace{\frac{1}{2} \Delta \mathcal{A} + \eta^{-1} \nabla \cdot (\hat{\mathbf{z}} \times \mathcal{F})}_{\stackrel{\text{def}}{=} q^w}, \quad (5.81)$$

where \mathcal{A} is the action

$$\mathcal{A} \stackrel{\text{def}}{=} \frac{1}{2f_0} |\phi|^2, \quad (5.82)$$

and \mathcal{F} is the action flux

$$\mathcal{F} \stackrel{\text{def}}{=} i \frac{\lambda^2}{4} (\phi \nabla \phi^* - \phi^* \nabla \phi). \quad (5.83)$$

The differential operators in (5.81) are

$$\Delta \stackrel{\text{def}}{=} \partial_x^2 + \partial_y^2 \quad \text{and} \quad \nabla \stackrel{\text{def}}{=} \partial_x \hat{\mathbf{x}} + \partial_y \hat{\mathbf{y}}. \quad (5.84)$$

Appendix 5.B shows that the action flux \mathcal{F} is parallel to the horizontal component of the “pseudomomentum” \mathbf{p}_h [Andrews and McIntyre 1978]:

$$\mathcal{F} = \eta \mathbf{p}_h. \quad (5.85)$$

Using (5.85) and standard vector identities yields an alternative expression for the wave PV:

$$q^w = \frac{1}{2} \Delta \mathcal{A} - \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{p}_h). \quad (5.86)$$

In the WKB limit, the wave PV is given by the curl of the pseudomomentum [Bühler and McIntyre 1998].

The two evolution equations are the dissipative YBJ equation

$$\phi_t + J(\psi^L, \phi) + \frac{1}{2}\zeta\phi - \frac{1}{2}\eta\Delta\phi = -\gamma\phi, \quad (5.87)$$

where we recall that $\gamma = \mu + m^2\nu$, and the balanced PV equation

$$q_t + J(\psi^L, q) = D_q, \quad (5.88)$$

with

$$D_q \stackrel{\text{def}}{=} -\mu\Delta\psi^E - 2\gamma mw_1 \quad (5.89)$$

$$= -\mu\zeta + (\gamma - \mu)\Delta\mathcal{A} - 2\frac{\gamma}{\eta}\nabla\cdot(\hat{\mathbf{z}} \times \mathcal{F}) \quad (5.90)$$

$$= -\mu\zeta + (2\gamma - \mu)\Delta\mathcal{A} - 2\gamma q^w. \quad (5.91)$$

The non-obvious D_q above is the upshot of the asymptotic analysis. Defining the horizontal average $\langle \rangle$, an important property is that $\langle D_q \rangle = 0$, so the bulk PV $\langle q \rangle$ is conserved in the dissipative NIW-QG system. But q can be locally generated by dissipation, as we now illustrate with a simple solution. Remarkably, the diffusivity κ does not appear in D_q .

5.4 Local PV generation—an illustrative example

To gain insight into the PV generation by wave dissipation, we consider an initial-value problem of Bretherton flow with $q = 0$ [Bühler and McIntyre 1998] associated with a Gaussian wave packet

$$\phi(x, y, 0) = \Phi \exp\left[-\frac{r^2}{2a^2} + i(kx + ly)\right], \quad (5.92)$$

where $r^2 \stackrel{\text{def}}{=} x^2 + y^2$ is the radial distance about the center of the packet, and Φ is a constant. Zero PV implies that ϕ determines the initial ($t = 0$) Lagrangian-mean flow through the wave PV q^w :

$$\Delta\psi^L = -q^w. \quad (5.93)$$

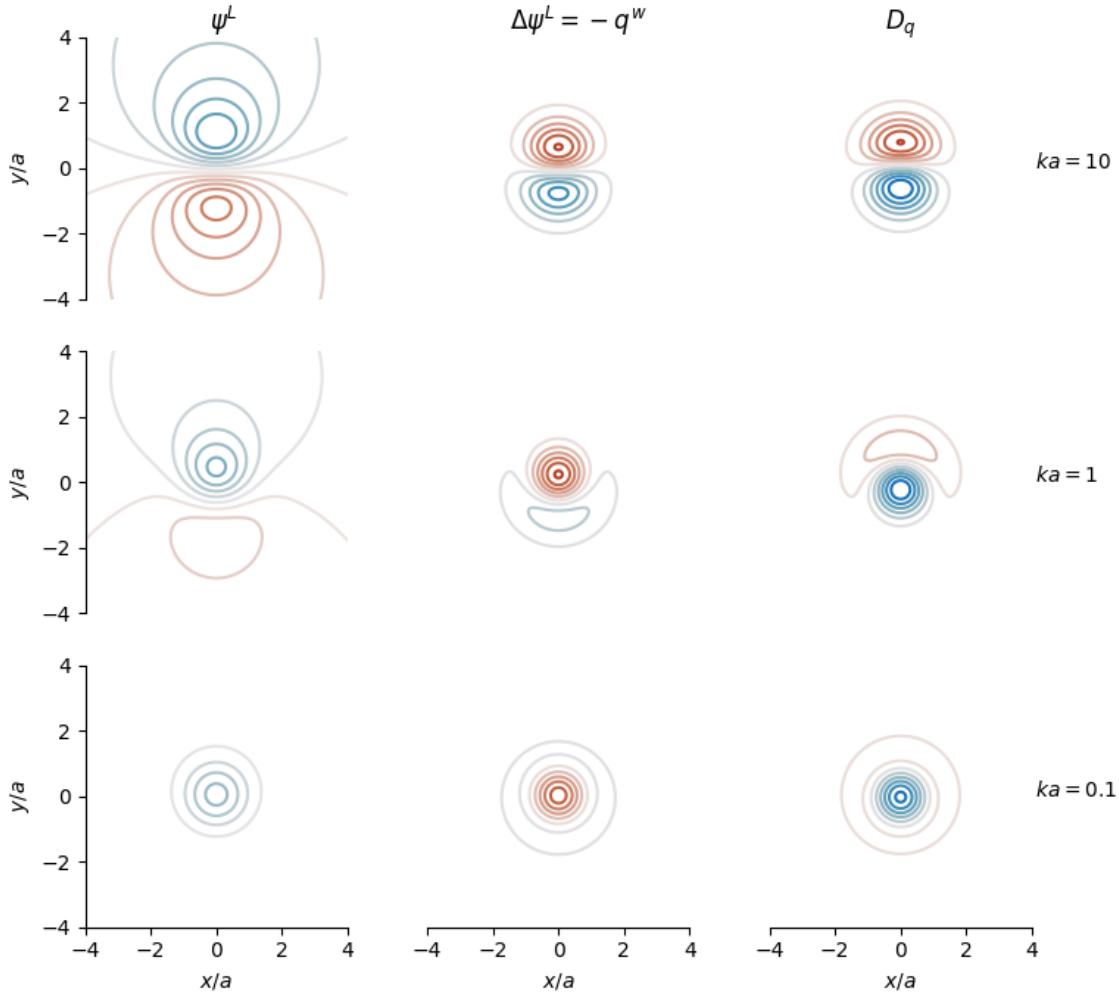


Figure 5.1: Lagrangian-mean streamfunction ψ^L in (5.99), wave PV q^w in (5.97), and wave dissipation in (5.98) for Gaussian wave packets (5.92) with $ka = (0.1, 1, 10)$ and $l = 0$.

The action density of the Gaussian packet (5.92) is an axis-symmetric envelope:

$$\mathcal{A} = \frac{|\Phi|^2}{2f_0} e^{-r^2/a^2}. \quad (5.94)$$

With the intermediate results

$$\Delta\mathcal{A} = \frac{4}{a^2} \left(\frac{r^2 - a^2}{a^2} \right) \mathcal{A} \quad (5.95)$$

and

$$\frac{2}{\eta} \nabla \cdot (\hat{\mathbf{z}} \times \mathcal{F}) = \frac{i}{f_0} J(\phi^*, \phi) = \frac{4}{a^2} (lx - ky) \mathcal{A}, \quad (5.96)$$

the wave PV is

$$q^w = \frac{2}{a^2} \left(lx - ky + \frac{r^2 - a^2}{a^2} \right) \mathcal{A}. \quad (5.97)$$

With $\mu = 0$, so that dissipation is only given by vertical viscosity $\gamma = m^2 \nu$, the PV dissipation associated with the wave packet (5.92) is

$$D_q = \frac{4\gamma}{a^2} \left(ky - lx + \frac{r^2 - a^2}{a^2} \right) \mathcal{A}. \quad (5.98)$$

On an infinite horizontal plane, we can invert the Laplacian in (5.93) to calculate the streamfunction ψ^L :

$$\psi^L = \frac{|\Phi|^2}{4f_0} \left[\frac{ky}{r^2} \left(e^{-r^2/a^2} - 1 \right) - e^{-r^2/a^2} \right]. \quad (5.99)$$

5.4.1 Two limiting solutions

Figure 5.1 depicts the streamfunction ψ^L , wave PV q^w , and wave dissipation D_q for a Gaussian wave packet with $l = 0$ and varying ka . There are two interesting limits of the solution (5.97)-(5.99). If $ka \gg 1$, the Lagrangian-mean streamfunction, wave PV, and wave dissipation reduce to

$$\psi^L \sim \frac{|\Phi|^2}{4f_0} \frac{ky}{r^2} \left(e^{-r^2/a^2} - 1 \right), \quad q^w \sim -\frac{2}{a^2} ky \mathcal{A}, \quad \text{and} \quad D_q \sim \frac{4\gamma}{a^2} ky \mathcal{A}, \quad (5.100)$$

with \mathcal{A} in (5.94) (see top panels of figure 5.1). This is the WKB limit in which the flow is a dipole [e.g., Bühler and McIntyre 1998]. The dissipative tendency D_q also has a dipolar spatial structure, so wave dissipation generates a q -dipole. Because D_q and q^w have opposite signs, wave dissipation intensifies the existing Lagrangian-mean flow ψ^L .

The opposite limit is $ka \ll 1$. In this inverse WKB limit, the solution reduces to

$$\psi^L \sim -\frac{1}{2} \mathcal{A}, \quad q^w \sim \frac{2}{a^2} \frac{r^2 - a^2}{a^2} \mathcal{A}, \quad \text{and} \quad D_q \sim \frac{4\gamma}{a^2} \frac{r^2 - a^2}{a^2} \mathcal{A}. \quad (5.101)$$

Because the action \mathcal{A} in (5.94) is a Gaussian, the solution (5.101) is a monopole. Both wave PV q^w and dissipative tendency D_q have the same bullseye spatial structure (see lower panels of figure 5.1). Contrary to the WKB limit, however, q^w and D_q have the same sign, so wave dissipation initially weakens the existing Lagrangian-mean flow.

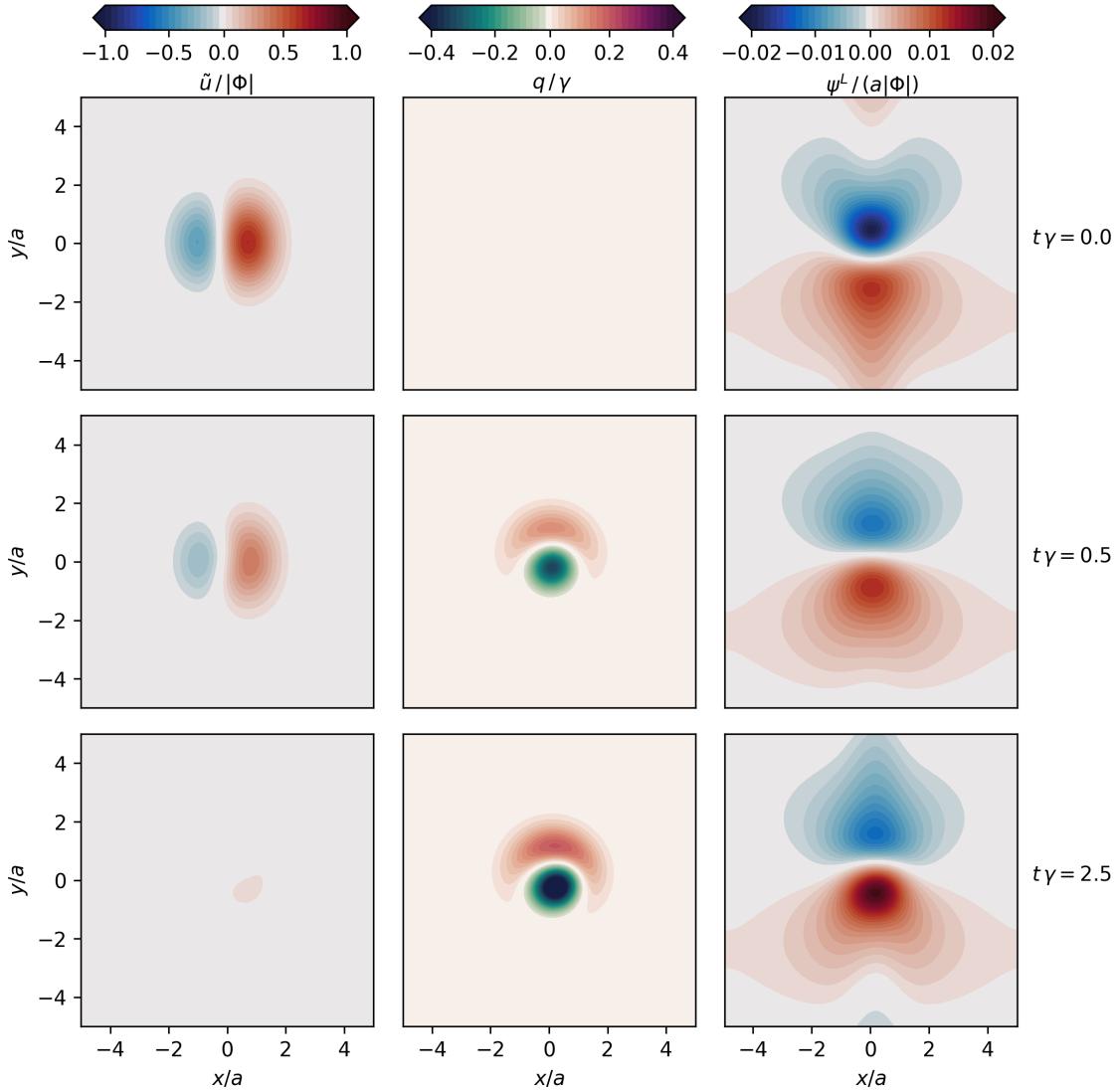


Figure 5.2: Solution of the dissipative NIW-QG model (5.87)-(5.88) with initial conditions $q = 0$ and the Gaussian wave packet (5.92) with $ka = 1$ and $l = 0$.

5.4.2 Nonlinear evolution on a periodic domain

The intermediate case $ka = 1$ is a superposition of these bullseye and dipolar solutions. To further illustrate the generation of PV by wave dissipation, we consider the time evolution of a $ka = 1$ wave packet on a periodic domain. Given the initial condition (5.92) and $q = 0$, we evolve the system (5.87)-(5.88) using a standard Fourier spectral method; for details about the numerical method, see Rocha, Wagner,

and Young [2018].

Figure (5.2) shows snapshots of the solution. Starting from zero PV everywhere, an asymmetric dipole is generated by wave dissipation, with stronger and more circular negative vorticity. The streamfunction of the initial condition is asymmetric, with stronger negative ψ^L . But wave dissipation generates a dipole with opposite sign, so ψ^L first weakens and become symmetric, and then enhances again, developing an asymmetry opposite to that of the initial condition.

These changes in the streamfunction ψ^L show that wave dissipation has non-monotonic effect on the Lagrangian-mean flow. To characterize the dissipative effects on the balanced flow, and to prepare the ground for the analysis of forced-dissipative solutions, we now turn into a detailed description of the energetics of the dissipative NIW-QG model in section (5.3.5).

5.5 Energetics of the dissipative QG-NIW model

As discussed in XV15 and Rocha, Wagner, and Young [2018], the inviscid vertical plane wave model conserves the domain-average of two quadratic quantities: wave action \mathcal{A} ,

$$\mathcal{A} \stackrel{\text{def}}{=} \frac{1}{2f_0} |\phi|^2, \quad (5.102)$$

and a coupled energy, the sum of geostrophic kinetic energy \mathcal{K}^L and wave potential energy \mathcal{P} ,

$$\mathcal{E} = \underbrace{\frac{1}{2} |\nabla \psi^L|^2}_{\stackrel{\text{def}}{=} \mathcal{K}^L} + \underbrace{\frac{\lambda^2}{4} |\nabla \phi|^2}_{\stackrel{\text{def}}{=} \mathcal{P}}. \quad (5.103)$$

The geostrophic kinetic energy \mathcal{K}^L is decorated with a superscript L because the Lagrangian-mean flow is in geostrophic balance through (5.46)-(5.47). \mathcal{K}^L is different than the kinetic energy of the Eulerian-mean flow. To understand the importance of this distinction, recall that the horizontal velocity is

$$\mathcal{U} = \mathcal{U}_0 + \epsilon \mathcal{U}_1 + \epsilon^2 \mathcal{U}_2 + \dots, \quad (5.104)$$

so the kinetic energy density is

$$\frac{1}{2} \overline{\mathcal{U}^* \mathcal{U}} = \frac{1}{2} \overline{\mathcal{U}_0^* \mathcal{U}_0} + \epsilon^2 \frac{1}{2} (\overline{\mathcal{U}_0^* \mathcal{U}_2} + \overline{\mathcal{U}_2^* \mathcal{U}_0} + \overline{\mathcal{U}_1^* \mathcal{U}_1}) + \mathcal{O}(\epsilon^3), \quad (5.105)$$

where we recall that \mathcal{U}_1 is independent of fast time—see (5.51) and surrounding discussion. Using (5.37), (5.51), and (5.149), and noting that phase average of the cross-terms is zero because \mathcal{U}_0 and \mathcal{U}_2 are out of phase, yields

$$\frac{1}{2}\overline{\mathcal{U}^*\mathcal{U}} = \frac{1}{2}|\phi|^2 + \epsilon^2 \underbrace{\frac{1}{2}|\nabla\psi^E|^2}_{\stackrel{\text{def}}{=} \mathcal{K}^E} + \mathcal{O}(\epsilon^3). \quad (5.106)$$

Thus the second-order kinetic energy is not the geostrophic kinetic energy \mathcal{K}^L .

In the parent Boussinesq equations (5.36), the domain-average of the Eulerian-mean—not the Lagrangian-mean—kinetic and potential energies decay monotonically due to dissipation. Hence, a consistent dissipative asymptotic model must dissipate the Eulerian energy in (5.106). Below we show that the dissipative NIW-QG model in section 5.3.5 preserves this fundamental property of the parent Boussinesq equations.

5.5.1 Wave action

The derivation of the energy equations follows Rocha, Wagner, and Young [2018]. First, multiplying the wave equation (5.87) by ϕ^* and adding the complex conjugate yields an equation for the action density:

$$\mathcal{A}_t + J(\psi^L, \mathcal{A}) + \nabla \cdot \mathcal{F} = -2\gamma \mathcal{A}, \quad (5.107)$$

where the action flux in (5.83) and we recall that $\gamma = \mu + m^2\nu^2$. With the horizontal average $\langle \rangle$, and assuming no-boundary-flux conditions, the action $\langle \mathcal{A} \rangle$ decays exponentially due to dissipation.

5.5.2 Wave potential energy

Next, multiplying (5.87) by $-\lambda^2\Delta\phi^*/4$, adding the complex conjugate, and averaging yields an equation for $\langle \mathcal{P} \rangle$:

$$\partial_t \langle \mathcal{P} \rangle = \Gamma_r + \Gamma_a - 2\gamma \langle \mathcal{P} \rangle, \quad (5.108)$$

where the refractive and advective “conversion terms” are [Rocha, Wagner, and Young 2018]

$$\Gamma_r \stackrel{\text{def}}{=} \left\langle \frac{1}{2}\zeta \nabla \cdot \mathcal{F} \right\rangle \quad (5.109)$$

and

$$\Gamma_a \stackrel{\text{def}}{=} \frac{\lambda^2}{4} \langle \psi^L [J(\phi, \Delta\phi^*) + J(\phi^*, \Delta\phi)] \rangle. \quad (5.110)$$

As expected, wave potential energy—the total potential energy of the asymptotic model—decays monotonically due to the final dissipative term in (5.108).

5.5.3 Geostrophic kinetic energy

The equation for \mathcal{K}^L follows from multiplying (5.88) by $-\psi^L$ and averaging:

$$\partial_t \langle \mathcal{K}^L \rangle = \langle \psi^L q_t^w \rangle - 2\mu \langle \mathcal{K}^L \rangle + (\mu - \gamma) \langle \zeta \mathcal{A} \rangle + 2\frac{\gamma}{\eta} \langle \psi^L \nabla \cdot (\hat{\mathbf{z}} \times \mathcal{F}) \rangle. \quad (5.111)$$

To work on $\langle \psi^L q_t^w \rangle$ on the right of (5.111) we use the action equation (5.107) and Ehrenfest's theorem in equation B12 of Rocha, Wagner, and Young [2018]. Using the intermediate results

$$\langle \psi^L \frac{1}{2} \Delta \mathcal{A}_t \rangle = -\Gamma_r - \frac{1}{2} \langle \zeta J(\psi^L, \mathcal{A}) \rangle - \gamma \langle \zeta \mathcal{A} \rangle, \quad (5.112)$$

and

$$\langle \psi^L \nabla \cdot (\hat{\mathbf{z}} \times \mathcal{F}_t) \rangle = \eta \frac{1}{2} \langle \zeta J(\psi^L, \mathcal{A}) \rangle - \Gamma_a - 2\gamma \langle \psi^L \nabla \cdot (\hat{\mathbf{z}} \times \mathcal{F}) \rangle, \quad (5.113)$$

we obtain

$$\partial_t \langle \mathcal{K}^L \rangle = -\Gamma_r - \Gamma_a - 2\mu \langle \mathcal{K}^L \rangle + (\mu - 2\gamma) \langle \zeta \mathcal{A} \rangle. \quad (5.114)$$

The rightmost term in (5.114) is sign indefinite. If $\mu = 0$, in particular, then dissipation of wave action, $-2\gamma \mathcal{A}$, in anticyclones, $\zeta < 0$, is a source of \mathcal{K}^L . But this does not imply that the dissipative effects on the Eulerian energy (5.106) are sign indefinite.

5.5.4 Eulerian-mean kinetic energy

To show that dissipation monotonically damps the Eulerian-mean kinetic energy, we recall that $\psi^L = \psi^E - \mathcal{A}$, and thus

$$\mathcal{K}^L = \underbrace{\frac{1}{2} |\nabla \psi^E|^2}_{\stackrel{\text{def}}{=} \mathcal{K}^E} + \underbrace{\frac{1}{2} |\nabla \mathcal{A}|^2}_{\stackrel{\text{def}}{=} \mathcal{K}^S} - \nabla \psi^E \cdot \nabla \mathcal{A}. \quad (5.115)$$

The identity

$$\langle \mathcal{K}^S \rangle = \langle \mathcal{A} q \rangle \quad (5.116)$$

suggests a simple route to an equation for the “Stokes energy” \mathcal{K}^S : multiplying (5.88) by \mathcal{A} and (5.107) by q , and averaging, yields

$$\partial_t \langle \mathcal{K}^S \rangle = -\underbrace{\langle q \nabla \cdot \mathcal{F} \rangle}_{\stackrel{\text{def}}{=} \Gamma_S} - \mu \langle \zeta \mathcal{A} \rangle - (\gamma - \mu) \langle |\nabla \mathcal{A}|^2 \rangle - 2\gamma \langle \mathcal{K}^S \rangle, \quad (5.117)$$

Finally, subtracting (5.117) from (5.114) yields an equation for $\langle \mathcal{K}^E \rangle$:

$$\partial_t \langle \mathcal{K}^E \rangle = \Gamma_S - \Gamma_r - \Gamma_a - 2\mu \langle \mathcal{K}^E \rangle, \quad (5.118)$$

where we used (5.115) and (5.116). Rocha, Wagner, and Young [2018] remark that $\Gamma_S - \Gamma_r - \Gamma_a$ is the Reynolds shear production (for details, see appendix 5.C). Thus \mathcal{K}^E changes due shear-production and Rayleigh dissipation; \mathcal{K}^E is not affected by viscosity because $u_z^E = 0$, and horizontal viscosity does not enter at this order.

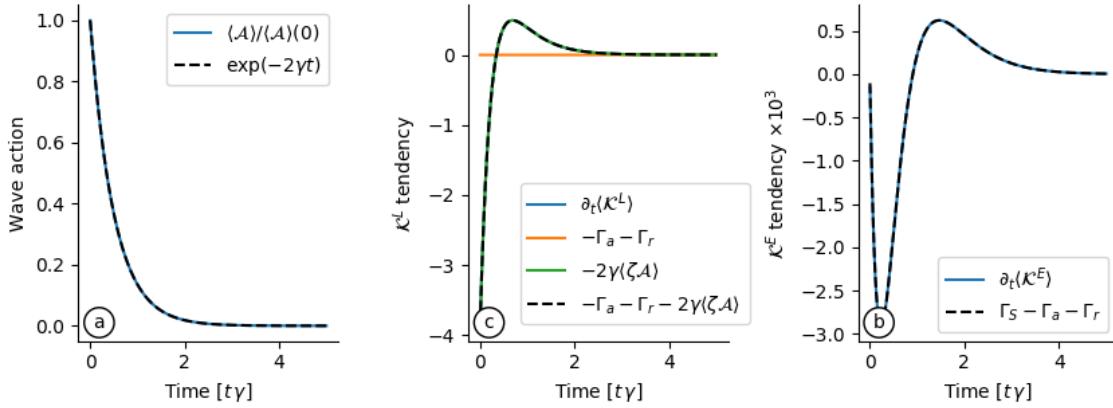


Figure 5.3: Energetics of a solution of the dissipative NIW-QG model (5.87)-(5.88) with initial conditions $q = 0$ and the Gaussian wave packet (5.92) with $ka = 1$ and $l = 0$. (a) Wave action, and (b) Lagrangian-mean and (c) Eulerian-mean kinetic energy budgets. The energy tendencies are normalized by $10^{-5} U_w^2 \gamma^{-1}$.

5.5.5 Energetics of the $ka = 1$ wave-packet solution

Figure 5.3 shows the energetics of the $ka = 1$ Gaussian-packet solution of figure 5.2. As expected, dissipation exponentially damps wave action. The Lagrangian-mean kinetic energy \mathcal{K}^L strongly decreases at very small time. This initial decrease in \mathcal{K}^L

is due PV dissipation $-2\gamma\langle\zeta\mathcal{A}\rangle$ —stimulated generation $\Gamma_a + \Gamma_r$ is 10^{-3} smaller than dissipation in this example—that creates a dipole with opposite sign of the existing one (see discussion surrounding figure 5.2). After $t\gamma \approx 0.5$, the \mathcal{K}^L tendency becomes positive, and this change in sign is associated with the enhancement of the dipole generated by PV dissipation. Because $\mu = 0$, the Eulerian-mean kinetic energy \mathcal{K}^E only changes due to shear production $\Gamma^S - \Gamma_r - \Gamma_a$ in (5.118), which is 10^{-3} smaller than the dissipative changes of \mathcal{K}^L .

5.6 Forced-dissipative solutions

With the rigorously derived dissipative model (5.3.5) and an understanding of the effects of wave dissipation on the balanced flow, we now turn into the investigation of stimulated generation in forced-dissipative solutions. The main goal here is to test the hypothesis that, without bottom drag, forced 2D turbulence equilibrates through stimulated generation. A secondary goal is to assess the extent that the insights about stimulated generation drawn from initial-value problems [e.g., Rocha, Wagner, and Young 2018]. hold in an equilibrated forced-dissipative system.

5.6.1 Forcing

We force the balanced turbulence by adding a stochastic forcing to the PV equation (5.88). The forcing is constant when $0 < t < \tau$, and then “renovates” and is constant in the second interval $\tau < t < 2\tau$. On the n th internal, $(n-1)\tau < t < n\tau$, the Fourier transform of this stochastic forcing is

$$\hat{F}_q = \sqrt{S_{|\mathbf{k}|}} \frac{\hat{\xi}_n}{\sqrt{\tau}}, \quad (5.119)$$

where $\hat{\xi}_n = \hat{\xi}_n^r + i\hat{\xi}_n^i$ is a complex normal random variable with variance σ_q^2 ,

$$\mathbb{E}[\xi_n] = 0 \quad \text{and} \quad \mathbb{E}[|\xi_n|^2] = \sigma_q^2, \quad (5.120)$$

\mathbb{E} denoting expectation. Also in (5.119), S is the spectrum of the forcing

$$S_{|\mathbf{k}|} = C(k^2 + l^2) \exp \left\{ -[(k^2 + l^2)^{1/2} - k_f]^2 / 2\Delta_f^2 \right\}, \quad (5.121)$$

where C is a constant determined by the normalization condition

$$\iint \frac{1}{2}C(k^2 + l^2)^{-1} S_{|k|} dk dl = 1. \quad (5.122)$$

Figure (5.4) shows the annular spectrum S and a realization of the forcing F_q , used below in the simulations of section 5.6.2. The white-noise limit requires a renovation time scale smaller than the numerical time step, $\tau \ll \Delta t$. Numerical implementation, however, forces us to choose $\tau = \Delta t$, so F_q is only an approximation to white noise. But the approximation is good enough because Δt is much smaller than any physical time scale of the QG-NIW problem. The main advantage of this white-noise forcing is that—without waves—the power delivered by the forcing is known [e.g., Srinivasan and Young 2012]:

$$\mathbb{E}[-\langle \psi F_q \rangle] = \sigma_q^2. \quad (5.123)$$

The numerical implementation of the forcing follows Srinivasan and Young [2012].

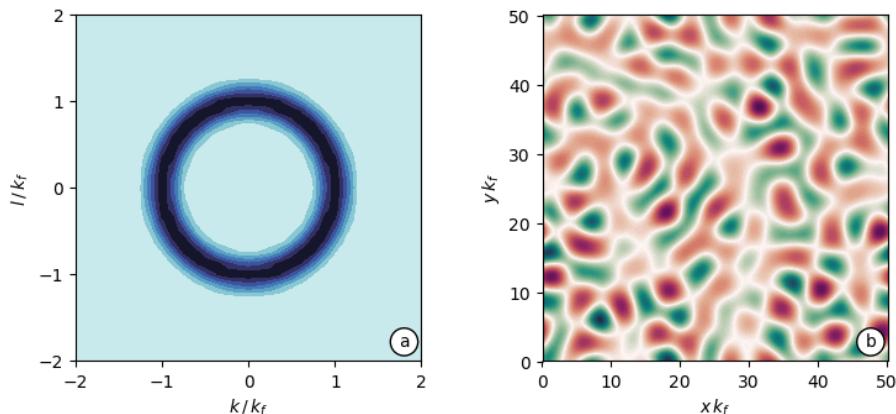


Figure 5.4: (a) Spectrum of the random PV forcing, an annulus centered at k_f with width Δ_f . (b) A realization of the random PV forcing.

The wave forcing added to the YBJ equation (5.87) is also an approximation to white-noise but with no spatial structure:

$$F_w = \frac{\xi_n}{\sqrt{\tau}}, \quad (5.124)$$

where $\xi_n = \xi_n^r + i\xi_n^i$ is a complex normal random variable with variance σ_w^2 ,

$$\mathbb{E}[\xi_n] = 0 \quad \text{and} \quad \mathbb{E}[|\xi_n|^2] = \sigma_w^2. \quad (5.125)$$

The motivation for this infinite-scale forcing is that atmospheric storms, which generate inertial currents, have lateral scales much larger than ocean eddies [c.f., Rocha, Wagner, and Young 2018]. The white-noise power delivered by F_ϕ is also known:

$$\mathbb{E}[\langle \phi^* F_w + \phi F_w^* \rangle] = \sigma_w^2. \quad (5.126)$$

The prediction for the work in (5.123) is the main motivation for forcing the waves with white noise. This type of forcing has been used previously both in idealized and Boussinesq numerical simulation [e.g., Taylor and Straub 2016; Barkan, Winters, and McWilliams 2016]. Tests with other types of forcing, such as shot noise and constant, yielded qualitatively similar results.

5.6.2 Three solutions

To illustrate the equilibration of forced 2D turbulence by stimulated generation, we numerically integrate the forced dissipative NIW-QG model in a periodic domain using standard Fourier spectral methods. To absorb the forward cascade of enstrophy q^2 and eliminate energy aliased wavenumbers, we use an exponential spectral filter when calculating the nonlinear terms in (5.81), (5.87), and (5.88). While important for grid-scale enstrophy mopping, the spectral filter has minimal impact on energy budgets because of the inverse cascade of balance kinetic energy and because there is no efficient forward cascade of wave energy in the vertical plane wave model [Rocha, Wagner, and Young 2018]. Table 5.1 contains all parameters of the *reference* NIW-QG solution. In particular,

$$\sigma_w = 4\sigma_q \quad \text{and} \quad m^2\nu = 3\mu, \quad (5.127)$$

so wave forcing and dissipation are stronger than eddy forcing and dissipation, and the equilibrated wave velocity is about twice the balance velocity $U_w \sim 2U_e$. This parameter regime is within the formal limits of validity of the NIW-QG model [see discussion in Rocha, Wagner, and Young 2018].

We also run two additional simulations, one without waves and the other without bottom drag. The parameters of the *no-wave* solution are same of the *reference* solution's in table 5.1 except that $\sigma_w = 0$; this solution is the classical problem of

Table 5.1: Parameters of the reference solution.

Parameter	Description	Value
L_d	Number of modes	512
σ_q^2	Domain size	$2\pi \times 200 \text{ km}$
σ_w^2	Balanced-forcing variance	$3.62 \cdot 10^{-9} \text{ m}^2 \text{ s}^{-3}$
$k_f L_d / 2\pi$	Wave-forcing variance	$5.78 \cdot 10^{-8} \text{ m}^2 \text{ s}^{-3}$
$dk_f L_d / 2\pi$	Balanced-forcing wavenumber	8
μ	Balanced-forcing width	1
ν	Linear bottom drag coefficient	$5.78 \cdot 10^{-8} \text{ s}^{-1}$
N	Viscosity	$1.98 \cdot 10^{-2} \text{ m}^2 \text{ s}^{-1}$
f_0	Buoyancy frequency	$5 \cdot 10^{-3} \text{ s}^{-1}$
$2\pi m^{-1}$	Coriolis frequency	$1 \cdot 10^{-4} \text{ s}^{-1}$
	Vertical wavelength	800 m

forced 2D turbulence and serves as the simplest test for the numerical code. The parameters of the *no-drag* solution are also in table 5.1 except that $\mu = 0$. This solution shows that forced 2D turbulence does equilibrate via stimulated generation without large-scale dissipation μ .

5.6.3 Equilibrated PV field

Before moving into a quantitative analysis of the solutions, it is instructive to look at qualitative aspects of the equilibrated flow. Figure 5.5 shows snapshots of PV of the three simulations. The snapshots show that the equilibrated q is similar across simulations, with ubiquitous features of 2D turbulence such as coherent vortices and filaments. A main difference is that the no-wave vorticity $q = \zeta$ is smoother than $q = \zeta + q^w$ of the wavy solutions; the latter is more filamented and irregular.

Despite the overall similarity between the q -fields of the reference and the no-drag solutions, the PV decomposition into relative vorticity ζ and q^w in (5.81) shows marked differences. The reference-solution ζ has large vortices, albeit more irregular than q -vortices. In the no-drag solution, however, the large q -vortices are not apparent in the relative vorticity ζ , which is instead dominated by small-scale “granulation”. Most of this small-scale structure in ζ cancel the small-scale wave PV q^w , yielding a much smoother q —the same cancellation is apparent in the reference solution. Rocha,

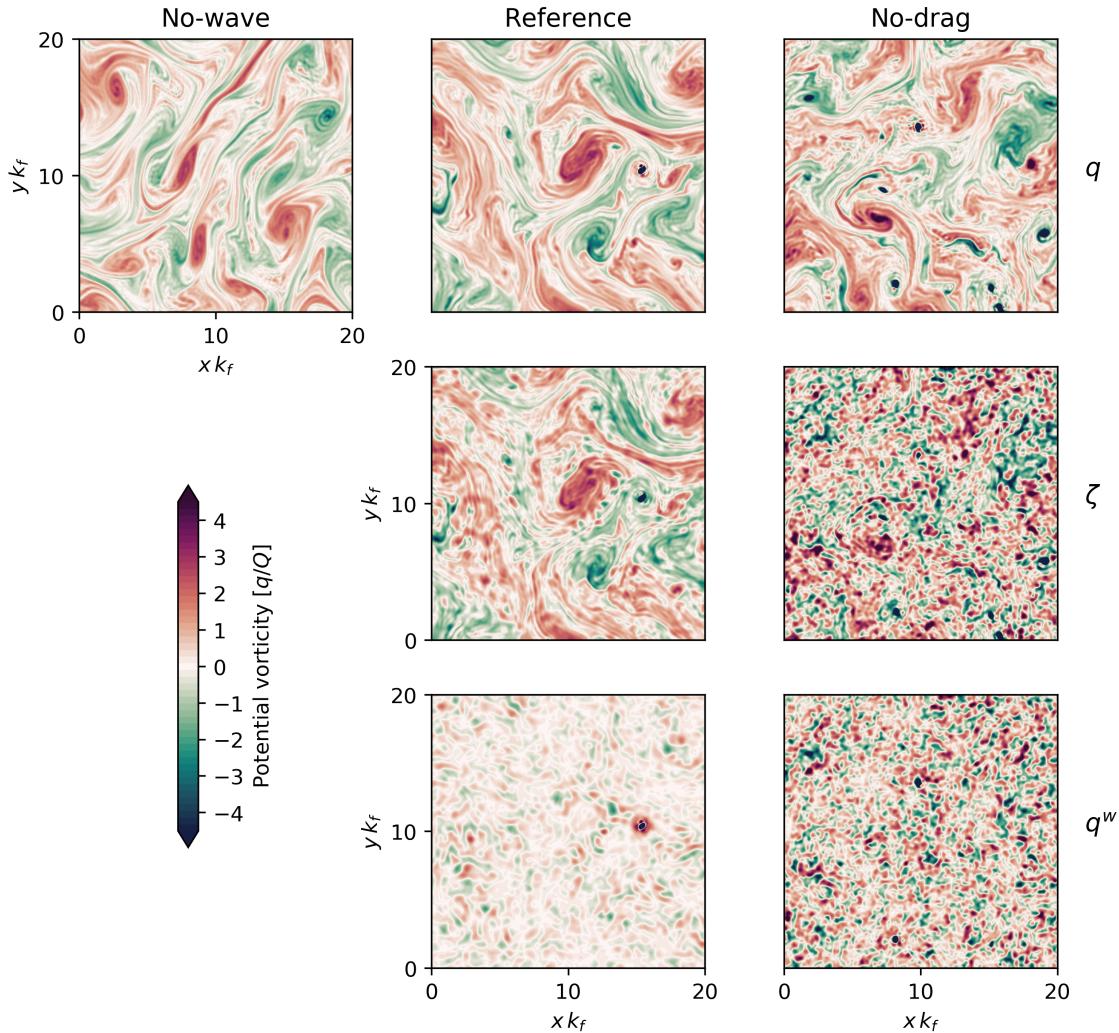


Figure 5.5: Snapshots at $t\gamma = 30$ of potential vorticity q and its decomposition into relative vorticity $\zeta \stackrel{\text{def}}{=} \Delta\psi^L$ and wave potential vorticity q^w defined in (5.81). (In the no-wave solution, $q^w = 0$ so that $q = \Delta\psi^L$.) The normalization constant is $Q = k_f U_e$ with $U_e = (\sigma_q^2/f_0)^{1/2}$. These plots show only $(1/2)^2$ of the simulation domain.

Wagner, and Young [2018] note that this ζ - q^w cancellation is a ubiquitous feature of NIW-QG solutions. And this cancellation underscores WY15’s remark that q in (5.81)—not the standard QG PV—is the most dynamically relevant quantity for phase-averaged systems such as the NIW-QG model in (5.3.5).

5.6.4 Energetics

Time series

Figure 5.6 shows time series of balanced kinetic energy \mathcal{K}^L , wave kinetic energy $f_0 \mathcal{A}$, and wave potential energy \mathcal{P} . The three solutions equilibrate in few dissipative time units γ^{-1} . As expected, \mathcal{K}^L in the no-wave solution equilibrates at the predicted level,

$$E_q \stackrel{\text{def}}{=} \frac{\sigma_q^2}{2\mu}, \quad (5.128)$$

with the parameters in table 5.1. The reference solution is less energetic than the no-wave solution, with \mathcal{K}^L oscillating about $0.6E_q$. The no-drag solution, however, is more energetic than the waveless prediction, with \mathcal{K}^L equilibrating at about $1.25E_q$.

The wave kinetic energy $f_0 \mathcal{A}$ shows small statistical differences between the reference and no-drag solutions. This insensitivity is expected because $\gamma = \mu + m^2\nu \approx m^2\nu$ and because the action budget in (5.107) is independent of the balanced flow. Thus wave kinetic energy equilibrates about the reference predicted level

$$E_w \stackrel{\text{def}}{=} \frac{\sigma_w^2}{2\gamma}, \quad (5.129)$$

with the parameters in table 5.1. In contrast to the balanced kinetic energy \mathcal{K}^L , wave kinetic energy $f_0 \mathcal{A}$ is intermittent, displaying fluctuations as large as $3E_w$.

Wave potential energy \mathcal{P} is significant in both wave solutions. Because wave forcing F_w in (5.4) has no spatial structure and therefore imparts only wave kinetic energy, stimulated generation $\Gamma_r + \Gamma_a$ is the only source of \mathcal{P} in (5.108). The resulting \mathcal{P} is significant in both solutions, with $0.07E_q$ for reference solution and $0.15E_q$ for no-drag solution. To quantify the role of stimulated generation in the equilibration of 2D turbulence, we now turn to the analysis of the balanced kinetic energy budgets.

Balanced kinetic energy budget

Figure 5.3 shows the \mathcal{K}^L budget in (5.114) time-averaged after the spin-up ($t\gamma \geq 10$). In the reference solution, half of the power imparted by the stochastic forcing is removed by bottom drag $-\mu\langle \mathcal{K}^L \rangle$; and the other half is converted into wave potential energy \mathcal{P} by stimulated generation, most of which is due to advective conversion Γ_a ;

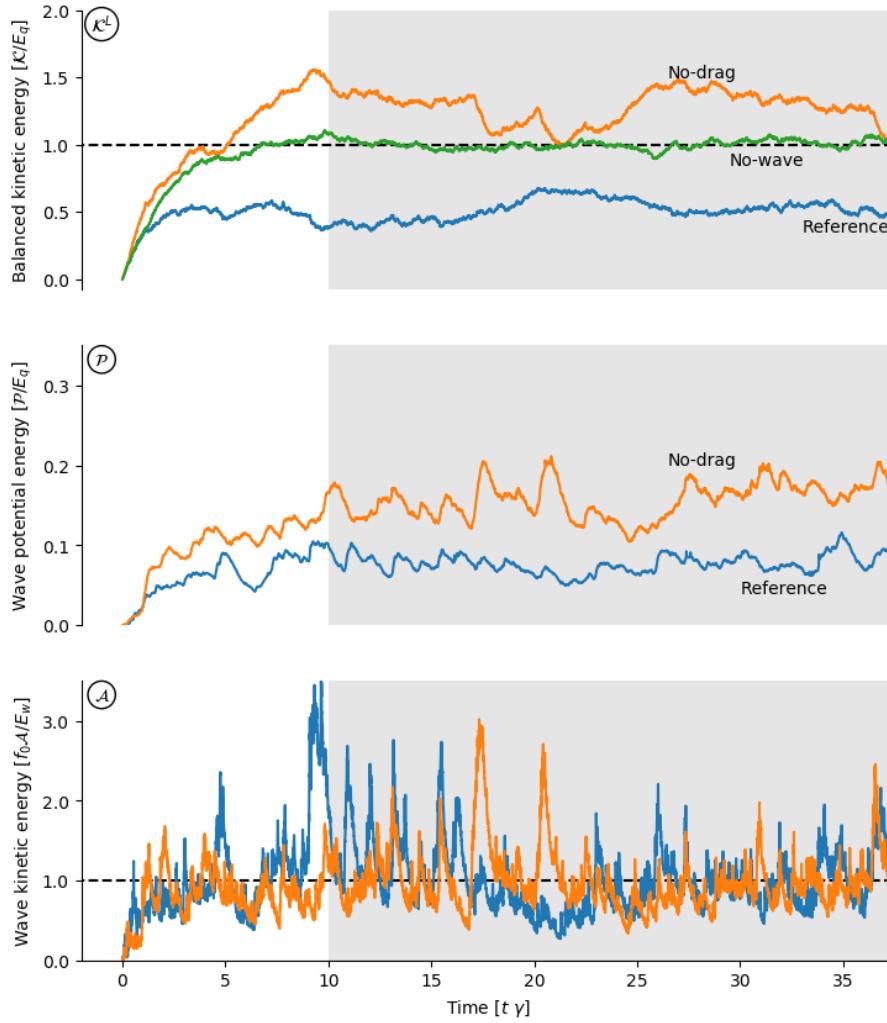


Figure 5.6: Time series of balanced kinetic energy \mathcal{K}^L , wave potential energy \mathcal{P} , and wave kinetic energy $f_0 \mathcal{A}$. \mathcal{K}^L and \mathcal{P} are normalized by the no-wave balanced kinetic energy prediction $E_q = \sigma_q^2/2\mu$, whereas $f_0 \mathcal{A}$ is normalized by the wave kinetic energy prediction $E_w = \sigma_w^2/2\gamma$, where $\gamma = \mu + m^2\nu$. Gray shading depicts the timespan used to compute averaged budgets.

the refractive conversion accounts for less than 20% of stimulated generation. Thus the main \mathcal{K}^L -balance of the reference solution is

$$\text{forcing} \approx \text{bottom drag} + \text{advective conversion}. \quad (5.130)$$

The nonlinear dissipative term $(\mu - 2\gamma)\langle \zeta \mathcal{A} \rangle \approx -2\gamma\langle \zeta \mathcal{A} \rangle$ in (5.114), associated with the generation of mean flow by wave dissipation, is non-negligible. This wave streaming source contributes about 12% of the total generation of balanced kinetic energy.

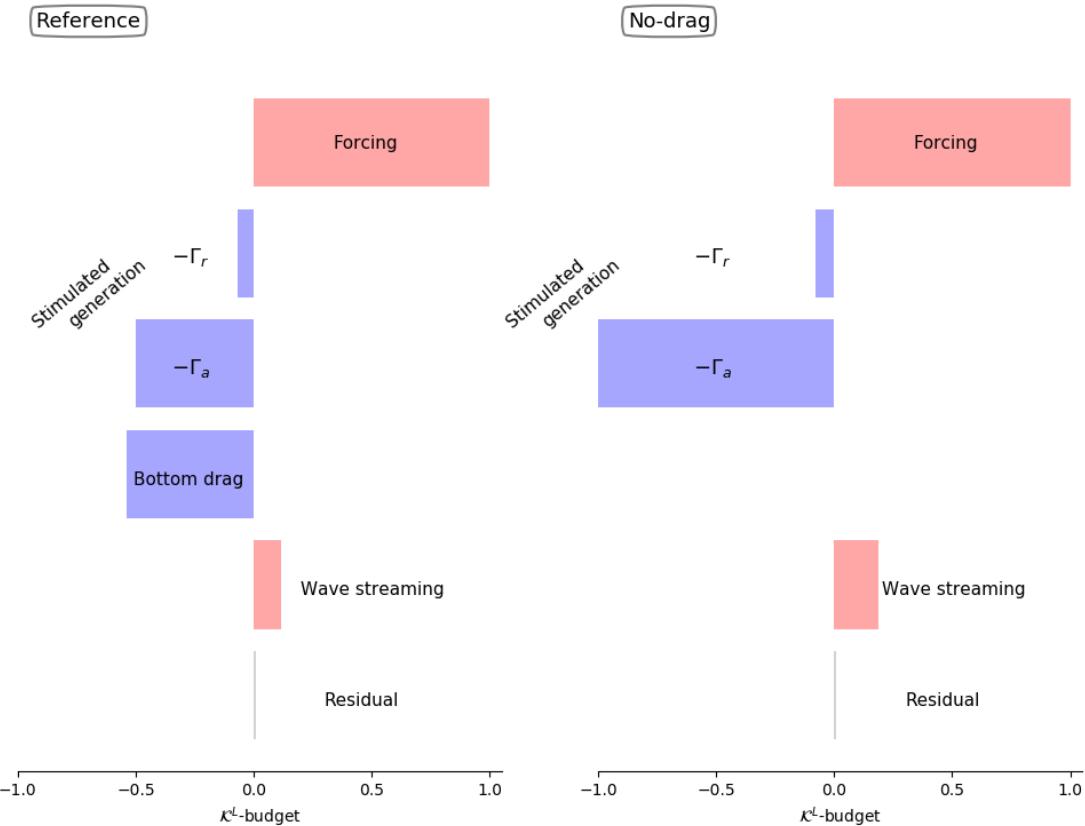


Figure 5.7: Balanced kinetic energy budget of reference and no-drag solutions. The budget represents an average over $t\gamma \geq 10$, and each term is normalized by the stochastic forcing. The residual is less than 1% in both budgets.

In the no-drag solution, most of the work delivered by stochastic forcing is removed by the advective conversion Γ_a . (The work delivered by the stochastic forcing has the same magnitude in both solutions.) In other words, most of the bottom-drag sink in the reference solution is replaced in the no-drag solution by a transfer of balanced kinetic energy \mathcal{K}^L into wave potential energy \mathcal{P} via advective conversion $-\Gamma_a$. Interestingly, wave streaming is stronger in the no-drag solution, contributing about 20% of the generation of \mathcal{K}^L .

In both wave solutions, about 95% of the wave kinetic energy $f_0 \mathcal{A}$ imparted by the forcing, σ_w^2 , is removed by dissipation $-2\gamma f_0 \langle \mathcal{A} \rangle$ in (5.107). Similarly, about 90% of the wave potential energy \mathcal{P} generated by stimulated generation is removed by $-2\gamma f_0 \langle \mathcal{P} \rangle$ in (5.108). Because $\gamma \approx m^2 \nu$, the bulk dissipation in the reference solution

stems from vertical viscosity in the parent Boussinesq equations (5.2). (In the no-drag solution, $\gamma = m^2 \nu$ so that all dissipation stems from vertical viscosity.) Small-scale dissipation by a spectral filter accounts for the residual in both wave energy budgets. As anticipated, this residual is small because wave escape prevents a forward cascade of wave energy [Rocha, Wagner, and Young 2018].

5.7 Conclusions

There are two main consequences of the dissipative NIW-QG model in section (5.3.5). First, the non-monotonic dissipative term D_q in (5.10) implies that wave dissipation can generate quasi-geostrophic PV q . And the analytical and numerical examples in section 5.4 extend and substantiate previous claims of PV generation by wave dissipation based on thought experiments that invoke “instant dissipation” in otherwise inviscid models [e.g., Bühler and McIntyre 1998]. Second, these non-monotonic dissipative effects on PV imply that wave dissipation can generate geostrophic, Lagrangian-mean kinetic energy \mathcal{K}^L in (5.114). This wave streaming mechanism has no effect on the Eulerian-mean kinetic energy \mathcal{K}^E in (5.118).

The forced-dissipative solutions in (5.6) confirm an important conclusion drawn from inviscid initial-value problems of stimulated generation [Rocha, Wagner, and Young 2018]: in the vertical plane wave model, geostrophic straining through Γ_a accounts for the bulk stimulated generation. Those solutions also show that through stimulated generation, forced 2D turbulence equilibrates even without bottom drag, consistent with the truncated primitive-equation double-gyre simulations of Gertz and Straub [2009]. A secondary result is that wave streaming, the rightmost term in (5.114), contributes a non-negligible source of geostrophic kinetic energy \mathcal{K}^L .

A growing number of Boussinesq numerical simulations—in which both balanced flow and waves are forced—show that externally forced internal waves enhance the dissipation of balanced flows [e.g., Taylor and Straub 2016; Barkan, Winters, and McWilliams 2016]. Within the theoretical framework of XV15 and WY16, the forced-dissipative NIW-QG system derived in section 5.3, and solutions such as those in section 5.6, may be useful in understanding these numerical simulations.

Appendix 5.A: Details about the model derivation

Why $\bar{U} = \epsilon \tilde{U}$?

To understand why $\bar{U} = \epsilon \tilde{U}$ is a convenient assumption, we decompose dynamical fields into mean plus waves, e.g., $u = \bar{u} + \tilde{u}$, where the overbar denotes Eulerian phase average with

$$\bar{\tilde{u}} = 0. \quad (5.131)$$

The phase-averaged potential vorticity is

$$\bar{\Pi} = f_0 N_0^2 + (\bar{v}_x - \bar{u}_y) N_0^2 + f_0 \bar{b}_z + \overline{\tilde{\omega} \cdot \nabla \tilde{b}} + \bar{\omega} \cdot \nabla \bar{b}. \quad (5.132)$$

With QG scaling,

$$\bar{v}_x - \bar{u}_y + \frac{f_0}{N_0^2} \bar{b}_z \sim \frac{\bar{U}}{L}, \quad (5.133)$$

and

$$\frac{\bar{\omega} \cdot \nabla \bar{b}}{N_0^2} \sim \underbrace{\frac{\bar{U}}{f_0 L}}_{\stackrel{\text{def}}{=} Ro} \frac{\bar{U}}{L}, \quad (5.134)$$

where Ro is the Rossby number. And with linear internal-wave scaling,

$$\frac{\overline{\tilde{\omega} \cdot \nabla \tilde{b}}}{N_0^2} \sim \underbrace{\frac{\tilde{U}}{f_0 L}}_{\stackrel{\text{def}}{=} \epsilon} \frac{\tilde{U}}{L}, \quad (5.135)$$

where ϵ is the wave non-linearity.

XV15 and WY15 identify the distinguished limit in which the non-linear wave potential vorticity terms enter *at the same order* of the QG potential vorticity. Comparing (5.135) with (5.133), this distinguished limit requires

$$\bar{U} = \epsilon \tilde{U}. \quad (5.136)$$

Appendix 5.B: Useful identities

Beltrami flow

The leading-order horizontal momentum equations are

$$u_{0t} - f_0 v = 0, \quad (5.137)$$

$$v_{0t} + f_0 u = 0, \quad (5.138)$$

and hence

$$\omega_{0t} = mw_z. \quad (5.139)$$

In the vertical plane wave model, the leading-order wave fields satisfy

$$m\theta_t + f\theta_z = 0. \quad (5.140)$$

Using (5.140) in (5.137)-(5.139) yields

$$mu_0 = -\omega_0. \quad (5.141)$$

Thus the leading-order motion is a Beltrami flow, with velocity parallel to vorticity.

Action flux and pseudomomentum

The hydrostatic horizontal component of the pseudomomentum of Andrews and McIntyre [1978] is (cf., WY15)

$$\mathbf{p}_h = -\overline{u_0 \nabla \xi_0} - \overline{v_0 \nabla \eta_0} - \frac{f_0}{2} (\overline{\xi_0 \nabla \eta_0} - \overline{\eta \nabla \xi_0}), \quad (5.142)$$

with the horizontal gradient operator $\nabla \stackrel{\text{def}}{=} \partial_x \hat{\mathbf{x}} + \partial_y \hat{\mathbf{y}}$. Noting that the horizontal near-inertial wave displacement is

$$f_0(\xi_0 + i\eta_0) = i \underbrace{\phi e^{i\omega}}_{=u_0 + iv_0}, \quad (5.143)$$

the pseudomomentum \mathbf{p}_h simplifies to

$$\mathbf{p}_h = \frac{f_0}{2} (\overline{\xi_0 \nabla \eta_0} - \overline{\eta \nabla \xi_0}). \quad (5.144)$$

Multiplying (5.143) by $\lambda^2 \nabla(5.143)^*/4$ and adding the complex conjugate yields an identity between the action flux \mathcal{F} and the pseudomomentum \mathbf{p}_h :

$$\mathcal{F} = \eta \mathbf{p}_h, \quad (5.145)$$

where we recall that $\eta = f_0 \lambda^2$ is the wave dispersivity. Thus, several \mathcal{F} interpretations in Rocha, Wagner, and Young [2018] are analogous to \mathbf{p}_h interpretations. For example, the refractive conversion $\Gamma_r = \langle \frac{1}{2} \zeta \nabla \cdot \mathcal{F} \rangle$ is positive when the pseudomomentum \mathbf{p}_h converges into anticyclones (and diverges from cyclones).

Appendix 5.C: Reynolds shear production

The Reynolds shear production is

$$RSP \stackrel{\text{def}}{=} \left\langle \bar{u}\bar{v}(\psi_{yy}^E - \psi_{xx}^E) + (\bar{u}^2 - \bar{v}^2)\psi_{xy}^E \right\rangle, \quad (5.146)$$

$$= i \left\langle \tilde{\mathcal{U}}^2 \psi_{ss}^E - \tilde{\mathcal{U}}^{*2} \psi_{s^*s^*}^E \right\rangle, \quad (5.147)$$

where ψ^E is the Eulerian-mean streamfunction and $\tilde{\mathcal{U}} = \bar{u} + i\bar{v}$ is the horizontal wave velocity, defined as the departure about the Eulerian-mean:

$$\tilde{\mathcal{U}} = \mathcal{U} - \mathcal{U}_1 = \mathcal{U}_0 + \mathcal{U}_2. \quad (5.148)$$

Also in (5.147) the Eulerian-mean streamfunction is $\psi^E = \psi^L + \mathcal{A}$. Using \mathcal{U}_0 in (5.37) and \mathcal{U}_2 in (5.149), we obtain

$$\overline{\tilde{\mathcal{U}}^2} = -2\phi \phi_2^*, \quad (5.149)$$

where we recall that

$$\phi_2 = f_0^{-1} \psi_{ss}^E \phi + \lambda^2 \phi_{ss}. \quad (5.150)$$

To evaluate RSP in terms of the conversions terms Γ 's of section (5.5), we introduce (5.149) into (5.147) and recall that $\psi^E = \mathcal{A} + \psi^L$. Using intermediate results such as

$$\langle \phi^* \phi_{ss} \psi_{s^*s^*}^L \rangle = \langle |\phi_s|^2 \psi_{ss^*}^L + \psi_{ss^*}^L \phi^* \phi_{ss^*} - \phi_s^* \phi_s \psi_{s^*s^*}^L \rangle, \quad (5.151)$$

yields

$$RSP = \Gamma_r - \Gamma_a + 2i\lambda^2 \langle \mathcal{A}_{ss^*} (\phi \phi_{ss^*}^* - \phi^* \phi_{ss^*}) \rangle + 2\lambda^2 \langle \mathcal{A}_{s^*s^*} i(\phi_s^* \phi_s - \mathcal{A}_{ss} \phi_{s^*} \phi_{s^*}^*) \rangle \quad (5.152)$$

Finally, using the identities (5.23) and (5.24), along with the results of section (5.5), gives an expression for RSP in terms of the conversionx terms Γ 's:

$$\begin{aligned} RSP &= \Gamma_r - \Gamma_a + \langle q^w \nabla \cdot \mathcal{F} \rangle, \\ &= \Gamma_s - \Gamma_r - \Gamma_a. \end{aligned} \quad (5.153)$$

Chapter 6

On Galerkin approximations of the surface-active quasigeostrophic equations

6.1 Introduction

Recent interest in upper-ocean dynamics and sub-mesoscale turbulence has focussed attention on surface geostrophic dynamics and the role of surface buoyancy variations. A main issue is the representation of active surface buoyancy by finite vertical truncations of the quasigeostrophic (QG) equations. Standard multi-layer [e.g., Pedlosky 1987] and modal approximations [e.g., Flierl 1978] assume that there is no variation of buoyancy on the surfaces.

Only few attempts have been made to represent both surface active and interior dynamics in the QG equations. The pioneering work by Tulloch and Smith [2009] developed a “two-mode two-surface” model that represents the surface dynamics exactly and approximates the interior dynamics using the barotropic and first baroclinic modes. The interaction of surface and interior dynamics motivated the development a new set of vertical modes that simultaneously diagonalize energy and a linear combination of enstrophy and surface buoyancy variance [Smith and Vanneste 2013]. Other studies of the interaction of surface and interior dynamics avoid vertical modes and use instead finite-difference schemes [Roullet et al. 2012] or idealize the interior potential vorticity

as a delta-function sheet [Callies et al. 2015].

Here we explore the representation of surface and interior dynamics using the familiar vertical modes of physical oceanography. These “standard modes”, denoted here by $p_n(z)$, are defined by the Sturm-Liouville eigenproblem

$$\frac{d}{dz} \frac{f_0^2}{N^2} \frac{dp_n}{dz} = -\kappa_n^2 p_n, \quad (6.1)$$

with homogeneous Neumann boundary conditions at the bottom ($z = z^-$) and top ($z = z^+$) surfaces of the domain:

$$\frac{dp_n}{dz}(z^\pm) = 0. \quad (6.2)$$

In (6.1) $N(z)$ is the buoyancy frequency and f_0 is the Coriolis parameter. The eigenvalue κ_n in (6.1) is the deformation wavenumber of the n 'th mode. With normalization, the modes satisfy the orthogonality condition

$$\frac{1}{h} \int_{z^-}^{z^+} p_n p_m dz = \delta_{mn}, \quad (6.3)$$

where $h \stackrel{\text{def}}{=} z^+ - z^-$ is the depth. The barotropic mode is $p_0 = 1$ and $\kappa_0 = 0$.

The modes defined by the eigenproblem (6.1) and (6.2) provide a fundamental basis for representing solutions of both the primitive and quasigeostrophic equations as a linear combination of $\{p_n\}$ [Gill 1982; Pedlosky 1987; Vallis 2006; Ferrari and Wunsch 2010; LaCasce 2012]. In fact, the set $\{p_n\}$ is mathematically complete and can be used to represent *any* field with finite square integral,

$$\int_{z^-}^{z^+} f^2 dz < \infty. \quad (6.4)$$

Even if the function $f(z)$ has nonzero derivative at z^\pm , or internal discontinuities, its representation as a linear combination of the basis functions $\{p_n\}$ converges in $L^2(z^-, z^+)$ i.e., the integral of the squared error goes to zero as the number of basis functions increases [e.g., Hunter and Nachtergael 2001, ch. 10]. In quasigeostrophic dynamics both the streamfunction ψ and the potential vorticity (PV) q satisfy the requirement (6.4) and thus both ψ and q can be effectively represented by linearly combining $\{p_n\}$.

Despite the rigorous assurance of completeness in the previous paragraph, the utility of $\{p_n\}$ for problems with nonuniform surface buoyancy has been questioned by

several authors [e.g., Lapeyre 2009; Roullet et al. 2012; Smith and Vanneste 2013]. These authors argue that the homogeneous boundary conditions in (6.2) are incompatible with nonzero surface buoyancy and that representation of the streamfunction ψ as a linear combination of $\{p_n\}$ is useless if ψ_z is nonzero on the surfaces.

The aim of this paper is to obtain a good Galerkin approximation to solutions of the QG equation with nonzero surface buoyancy using the familiar basis $\{p_n\}$. We show that both the inversion problem *and* evolutionary dynamics can be handled using $\{p_n\}$ to represent the streamfunction. As part of this program we revisit and extend two existing modal approximations [Flierl 1978; Tulloch and Smith 2009], and develop a new Galerkin approximation. We discuss the relative merit of the three approximations in terms of their mathematical rigor and conservation laws, and illustrate their efficiency and caveats by solving linear stability problems with nonzero surface buoyancy.

Using concrete examples, we show that the concerns expressed by earlier authors regarding the suitability of the standard modes $\{p_n\}$ are over-stated: even with nonzero surface buoyancy, the Galerkin expansion of the streamfunction ψ in terms of $\{p_n\}$ converges absolutely and uniformly with no Gibbs phenomena, and the same is true for the potential vorticity q . A modest number of terms provides a good approximation to ψ and q throughout the domain, including on the top and bottom boundaries. In other words, the surface streamfunction can be expanded in terms of $\{p_n\}$ and, with enough modes, this representation can then be used to accurately calculate the advection of nonzero surface buoyancy. In section 6.5 we illustrate this procedure by solving the classic Eady problem using the basis $\{p_n\}$ for the streamfunction.

6.2 The exact system

In this section we summarize the basic properties of the QG system. For a detailed derivation see Pedlosky [1987].

6.2.1 Formulation

The streamfunction is denoted $\psi(x, y, z, t)$ and we use the following notation.

$$u = -\psi_y, \quad v = \psi_x, \quad \vartheta = \left(\frac{f_0}{N}\right)^2 \psi_z. \quad (6.5)$$

The variable ϑ is related to the buoyancy by $b = N^2\vartheta/f_0$. The QG potential vorticity (QGPV) equation is

$$\partial_t q + J(\psi, q) + \beta v = 0, \quad (6.6)$$

where the potential vorticity is

$$q = (\Delta + L)\psi, \quad (6.7)$$

with

$$\Delta \stackrel{\text{def}}{=} \partial_x^2 + \partial_y^2, \quad \text{and} \quad L \stackrel{\text{def}}{=} \partial_z \left(\frac{f_0}{N} \right)^2 \partial_z. \quad (6.8)$$

Also in (6.6), the Jacobian is $J(A, B) \stackrel{\text{def}}{=} \partial_x A \partial_y B - \partial_y A \partial_x B$.

The boundary conditions at the top ($z = z^+$) and bottom ($z = z^-$) are that $w = 0$, or equivalently

$$@z = z^\pm : \quad \partial_t \vartheta^\pm + J(\psi^\pm, \vartheta^\pm) = 0. \quad (6.9)$$

Above we have used the superscripts $+$ and $-$ to denote evaluation at z^+ and z^- e.g., $\psi^+ = \psi(x, y, z^+, t)$.

6.2.2 Quadratic conservation laws

In the absence of sources and sinks, the exact QG system has four quadratic conservation laws: energy, potential enstrophy, and surface buoyancy variance at the two surfaces [e.g., Pedlosky 1987; Vallis 2006]. Throughout we assume horizontal periodic boundary conditions.

The well-known energy conservation law is

$$\frac{d}{dt} \underbrace{\int \frac{1}{2} |\nabla \psi|^2 + \frac{1}{2} \left(\frac{f_0}{N} \right)^2 (\partial_z \psi)^2 dV}_{\stackrel{\text{def}}{=} E} = 0. \quad (6.10)$$

The total energy is $\rho_0 E$, where ρ_0 is a reference density.

If $\beta = 0$ then there are many quadratic potential enstrophy invariants: the volume integral of $q^2 A(z)$, with $A(z)$ an arbitrary function of the vertical coordinate, is conserved. The choice $A(z) = \delta(z - z_*)$ reduces to conservation of the surface integral of q^2 at any level z_* .

Charney [1971] noted that, in a doubly periodic domain, nonzero β destroys all these quadratic potential enstrophy conservation laws, including the conservation of

potential enstrophy defined simply as the volume integral of q^2 . Multiplying the QGPV equation (6.6) by q , and integrating by parts, we obtain

$$\frac{d}{dt} \int \frac{1}{2} q^2 dV + \beta \int [v\vartheta]_{z^-}^{z^+} dS = 0. \quad (6.11)$$

The potential enstrophy equation (6.11) is the finite-depth analog of equation (13) in Charney's paper. To make progress Charney assumed $\vartheta = 0$ at the ground. But the β -term on the right of (6.11) can be eliminated by cross-multiplying the QGPV equation (6.6) evaluated at the surfaces z^\pm with the boundary conditions (6.9), and combining with (6.11). Thus, in a doubly periodic domain, nonzero β selects a uniquely conserved potential enstrophy from the infinitude of $\beta = 0$ potential enstrophy conservation laws:

$$\underbrace{\frac{d}{dt} \int \frac{1}{2} q^2 dV - \int q^+ \vartheta^+ - q^- \vartheta^- dS}_{\stackrel{\text{def}}{=} Z} = 0. \quad (6.12)$$

With $\beta \neq 0$ the surface contributions in (6.12) are required to form a conserved quadratic quantity involving q^2 . Notice that Z is not sign-definite. To our knowledge, the conservation law in (6.12) is previously unremarked.

Finally, in addition to E and Z , the surface buoyancy variance is conserved on each surface

$$\frac{d}{dt} \int \frac{1}{2} (\vartheta^\pm)^2 dS = 0. \quad (6.13)$$

Thus, with $\beta \neq 0$, the QG model has four quadratic conservation laws: E , Z and the buoyancy variance at the two surfaces.

6.3 Galerkin approximation using standard vertical modes

A straightforward approach is to represent the streamfunction by linearly combining the first $N + 1$ vertical modes. The mean square error in this approximation is

$$\text{err}_\psi(a_0, a_1, \dots, a_N) \stackrel{\text{def}}{=} \frac{1}{h} \int_{z^-}^{z^+} \left(\psi - \sum_{n=0}^N a_n p_n \right)^2 dz. \quad (6.14)$$

We use a roman font, and context, to distinguish the truncation index N in (6.14) from the buoyancy frequency $N(z)$. The coefficients a_0 through a_N are determined

to minimize err_ψ , and thus one obtains the Galerkin approximation ψ_N^G to the exact streamfunction:

$$\psi_N^G(x, y, z, t) \stackrel{\text{def}}{=} \sum_{n=0}^N \check{\psi}_n(x, y, t) p_n(z), \quad (6.15)$$

where the coefficients in the sum above are

$$\check{\psi}_n(x, y, t) \stackrel{\text{def}}{=} \frac{1}{h} \int_{z^-}^{z^+} \psi p_n dz. \quad (6.16)$$

Throughout we use the superscript $\check{\cdot}$ to denote a Galerkin coefficient defined via projection of a field onto a vertical mode.

In complete analogy with the streamfunction, one can also develop an $(N + 1)$ -mode Galerkin approximation to the PV:

$$q_N^G(x, y, z, t) \stackrel{\text{def}}{=} \sum_{n=0}^N \check{q}_n(x, y, t) p_n(z), \quad (6.17)$$

with coefficients

$$\check{q}_n \stackrel{\text{def}}{=} \frac{1}{h} \int_{z^-}^{z^+} q p_n dz. \quad (6.18)$$

The construction of the Galerkin approximation q_N^G above minimizes a mean square error err_q defined in analogy with (6.14).

Now recall that the exact ψ and q are related by the elliptic “inversion problem”:

$$(\Delta + L)\psi = q, \quad (6.19)$$

with boundary conditions at z^\pm :

$$\left(\frac{f_0}{N}\right)^2 \psi_z = \vartheta^\pm. \quad (6.20)$$

The Galerkin approximations in (6.15) through (6.18) are defined independently of the information in (6.19) and (6.20). The relationship between the Galerkin coefficients \check{q}_n and $\check{\psi}_n$ is obtained by multiplying (6.19) by $\frac{1}{h} p_n(z)$ and integrating over the depth. Noting the intermediate result

$$\frac{1}{h} \int_{z^-}^{z^+} p_n L \psi dz = \frac{1}{h} [p_n^+ \vartheta^+ - p_n^- \vartheta^-] - \kappa_n^2 \check{\psi}_n, \quad (6.21)$$

we obtain

$$\check{q}_n = \Delta_n \check{\psi}_n + \underbrace{\frac{1}{h} (p_n^+ \vartheta^+ - p_n^- \vartheta^-)}_{\text{surface terms}}, \quad (6.22)$$

where Δ_n is the n 'th mode Helmholtz operator

$$\Delta_n \stackrel{\text{def}}{=} \Delta - \kappa_n^2. \quad (6.23)$$

The relation in (6.22) is the key to a good Galerkin approximation to surface-active quasigeostrophic dynamics.

Term-by-term differentiation of the ψ_N^G -series in (6.15) does not give the q_N^G series in (6.17) unless $\vartheta^\pm = 0$. In other words, term-by-term differentiation does not produce the correct relation (6.22) between \check{q}_n and $\check{\psi}_n$. Thus the Galerkin truncated PV and the Galerkin truncated streamfunction do not satisfy the inversion boundary value problem exactly

$$(\Delta + L) \psi_N^G \neq q_N^G. \quad (6.24)$$

Despite (6.24), the truncated series ψ_N^G and q_N^G are the best least-squares approximations to ψ and q .

Notice that, in analogy with the Galerkin approximations for q and ψ ,

$$\check{\delta}_n^+ = \frac{1}{h} p_n^+ \quad \text{and} \quad \check{\delta}_n^- = \frac{1}{h} p_n^-, \quad (6.25)$$

where

$$\delta_N^{+G}(z) = \sum_{n=0}^N \check{\delta}_n^+ p_n \quad \text{and} \quad \delta_N^{-G}(z) = \sum_{n=0}^N \check{\delta}_n^- p_n, \quad (6.26)$$

are finite approximations to distributions $\delta(z - z^\pm)$ at the surfaces. Of course, these surface δ -distributions do not satisfy the L^2 convergence condition in (6.4) and thus the series in (6.26) only converge in a distributional sense [e.g., Hunter and Nachtergael 2001]. For instance, if f satisfies the L^2 convergence condition in (6.4), then

$$\int_{z^-}^{z^+} f(z) \delta_N^{+G}(z) dz \rightarrow \int_{z^-}^{z^+} f(z) \delta(z - z^+) dz = f(z^+), \quad (6.27)$$

as $N \rightarrow \infty$. Thus, in that limit,

$$(\Delta + L) \psi_N^G \rightarrow q - \delta(z - z^+) \vartheta^+ + \delta(z - z^-) \vartheta^-, \quad (6.28)$$

where \rightharpoonup denotes distributional convergence. The right-hand-side of (6.28) is the Brethertonian modified potential vorticity [Bretherton 1966] with the boundary conditions incorporated as PV sheets. To illustrate (6.24) and (6.28) we present an elementary example that is relevant to our discussion of the Eady problem in section 6.5.

An elementary example: the Eady basic state

As an example, consider the case with constant buoyancy frequency N . We use nondimensional units so that the surfaces are at $z^- = -1$ and $z^+ = 0$. The standard vertical modes are $p_0 = 1$ and, for $n \geq 1$

$$p_n = \sqrt{2} \cos(n\pi z), \quad (6.29)$$

with $\kappa_n = n\pi$.

We consider the basic state of the Eady problem with streamfunction

$$\psi = -\underbrace{(1+z)}_U y, \quad (6.30)$$

and zero interior PV $q = 0$ and $\beta = 0$. The surface buoyancies are $\vartheta^\pm = -y$.

The Galerkin expansion of the PV $q = 0$ is exact: $\check{q}_N = 0$ and therefore $q_N^G = 0$. The truncated Galerkin expansion of ψ follows from either (6.16) or (6.22) and is

$$\psi_N^G = -\underbrace{\left[\frac{1}{2}p_0 + 2\sqrt{2} \left(\frac{p_1}{\pi^2} + \frac{p_3}{(3\pi)^2} + \cdots + \frac{p_N}{(N\pi)^2} \right) \right]}_{U_N^G} y. \quad (6.31)$$

(We assume that N is odd, so that the last term in the truncated series is as above.) Despite the nonzero derivative of ψ at the boundaries, the series in (6.31) is absolutely and uniformly convergent on the closed interval $-1 \leq z \leq 0$. The N^{-2} behavior of the series (6.31) ensures uniform convergence, e.g., using the M-test [Hunter and Nachtergael 2001]. There are no Gibbs oscillations and a modest number of terms provides a good approximation to the base velocity U (Figure 6.1a). All these desirable properties are lost if we differentiate (6.31) with respect to z .

Thus, to illustrate (6.24) and (6.28), notice that if one attempts to calculate the Eady PV, namely $q = 0$, by direct differentiation of (6.31), one obtains

$$(\Delta + L) \psi_N^G = 2\sqrt{2}(p_1 + p_3 + \dots + p_N) y \quad (6.32)$$

$$= 2 \frac{\sin[(N+1)\pi z]}{\sin(\pi z)} y. \quad (6.33)$$

The series (6.32) does not converge in a pointwise sense and the partial sum is violently oscillatory as $N \rightarrow \infty$. However, in a distributional sense Hunter and Nachtergael [2001, ch. 11], the exact sum in (6.33) does converge to δ -distributions on the boundaries; see figures 6.1(b) and 6.1(c). These boundary δ -distributions are the Brethertonian PV sheets [Bretherton 1966]. To some extent, which we investigate in section 5(a), the series (6.32) is rescued by this Brethertonian interpretation.

Of course the correct Galerkin approximation to the Eady PV $q = 0$ is the splendidly convergent series $0 = 0p_0 + 0p_1 + \dots$, which is obtained if one uses either (6.18) or (6.22) to obtain $\check{q}_n = 0$. This seemingly trivial example illustrates potentially confusing issues that arise if one differentiates a Galerkin approximation: the standard modes provide good representations of ψ and q , even if ψ_z is nonzero on the boundaries. The problem is that differentiating the ψ -series to produce a q -series does not produce the Galerkin approximation to q .

6.4 Three approximations

In (6.24) we noted that the Galerkin approximations to ψ and q do not exactly satisfy the inversion relation. To address this error there are at least three different approximations one can make. The three approximations are equivalent when $\vartheta^\pm = 0$. In the next three sub-sections, we provide a detail description of each approximation. After testing, we recommend approximation C as the most reliable approximation using standard vertical modes.

6.4.1 Approximation A

Approximation A uses the truncated series ψ_N^G in (6.15) as a least-squares Galerkin approximation to the streamfunction ψ . A does not use the Galerkin approxi-

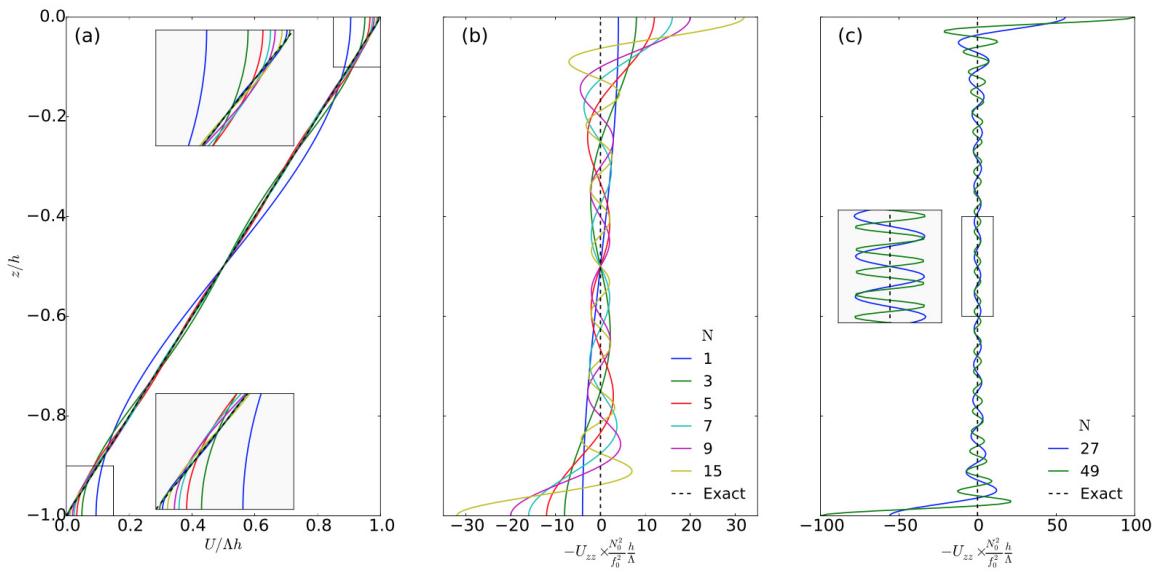


Figure 6.1: Nondimensional base-state for the Eady problem using various truncation for the series (6.31). In the middle panel N is the number of baroclinic modes. (a) Zonal velocity: although the truncation has zero slope at the boundaries there are no Gibbs oscillations. (b) Meridional PV gradient associated with the truncated series (6.33). (c) as in (b) but with an expanded abscissa. As N increases, the PV gradient distributionally converges to two Brethertonian delta functions at the boundaries.

mation for q . Instead, the approximate PV, $q_N^A(x, y, z, t)$, is *defined* so that the interior inversion relation is satisfied exactly:

$$q_N^A \stackrel{\text{def}}{=} (\Delta + L) \psi_N^G. \quad (6.34)$$

This is the approximation introduced by Flierl [1978] in a context without surface buoyancy, and it is now regarded as the standard in physical oceanography. Note that q_N^A in (6.34) is not the least-squares approximation to the exact q . Instead the series q_N^A is obtained using term-by-term differentiation of the series ψ_N^G . The example surrounding (6.32) shows that with nonzero surface buoyancy, the approximation q_N^A is strongly oscillatory in the interior of the domain and approaches the Brethertonian PV on the right of (6.28) as $N \rightarrow \infty$. The rapid interior oscillation of q_N^A is a spurious creation of term-by-term differentiation. Later, in section 6.5, these spurious oscillations will produce significant errors in the solution of the Eady stability problem.

Following Flierl [1978], in approximation A the N-mode approximate PV is defined via (6.34) and, using the modal representation for ψ_N^G in (6.15), this is equivalent to

$$q_N^A \stackrel{\text{def}}{=} \sum_{n=0}^N \Delta_n \check{\psi}_n(x, y, t) p_n(z), \quad (6.35)$$

where Δ_n is the Helmholtz operator in (6.23). Following the appendix of Flierl [1978], one can use Galerkin projection of the nonlinear evolution equation (6.6) onto the modes p_n to obtain $N + 1$ evolution equations for the coefficients $\check{\psi}_n$:

$$\partial_t \Delta_n \check{\psi}_n + \sum_{m=0}^N \sum_{s=0}^N \Xi_{nms} J(\check{\psi}_m, \Delta_s \check{\psi}_s) + \beta \partial_x \check{\psi}_n = 0, \quad (6.36)$$

where

$$\Xi_{nms} \stackrel{\text{def}}{=} \frac{1}{\bar{h}} \int_{z^-}^{z^+} p_n p_m p_s dz. \quad (6.37)$$

Note that Ξ_{nms} cannot be computed exactly except in cases with simple buoyancy frequency profiles. But it suffices to compute Ξ_{nms} to high accuracy, e.g. using Gaussian quadrature.

Flierl [1978] implicitly assumed that $\vartheta^+ = \vartheta^- = 0$, so that the surface terms in (6.22) vanish and then there is no difference between q_N^A and q_N^G . But in general, with nonzero surface buoyancy, we can append evolution equations for ϑ^+ and ϑ^- to

approximation A. That is, in addition to the $N + 1$ modal equations in (6.36), we also have

$$\partial_t \vartheta^\pm + \sum_{n=0}^N p_n^\pm J(\check{\psi}_n, \vartheta^\pm) = 0. \quad (6.38)$$

Above we have evaluated the ψ -series (6.15) at z^\pm to approximate ψ^\pm in the surface boundary conditions. This approach is not satisfactory because the resulting surface buoyancy equations (6.38) are dynamically passive i.e., ϑ^+ and ϑ^- do not affect the interior evolution equations in (6.36).

Approximation A has the well-known energy conservation law

$$\frac{d}{dt} \sum_{n=0}^N \int \frac{1}{2} (\nabla \check{\psi}_n)^2 + \frac{1}{2} \kappa_n^2 \check{\psi}_n^2 dS = 0. \quad (6.39)$$

To obtain the energy analogous to E in (6.10), the modal sum above is multiplied by the depth h . Approximation A also has the potential enstrophy conservation law,

$$\frac{d}{dt} \sum_{n=0}^N \int \frac{1}{2} (\Delta_n \check{\psi}_n)^2 dS = 0. \quad (6.40)$$

But the analog of the exact potential enstrophy (6.12) is not conserved by A (nor by B and C below). Finally, with the surface equations in (6.38), approximation A also conserves surface buoyancy variance as in (6.13).

6.4.2 Approximation B

Approximation B begins with the observation that the exact solution of the inversion problem in (6.19) and (6.20) can be decomposed as

$$\psi = \psi^I + \psi^S \quad (6.41)$$

where $\psi^I(x, y, z, t)$ is the “interior streamfunction” and $\psi^S(x, y, z, t)$ is the “surface streamfunction” [Lapeyre and Klein 2006; Tulloch and Smith 2009].

The surface streamfunction $\psi^S(x, y, z, t)$ is defined as the solution of the boundary value problem

$$(\Delta + L) \psi^S = 0, \quad (6.42)$$

with inhomogeneous Neumann boundary conditions

$$\left(\frac{f_0}{N}\right)^2 \partial_z \psi^S(z^\pm) = \vartheta^\pm. \quad (6.43)$$

The interior streamfunction $\psi^I(x, y, z, t)$ is defined as the solution of the boundary value problem

$$(\Delta + L) \psi^I = q, \quad (6.44)$$

with homogeneous Neumann boundary conditions

$$\left(\frac{f_0}{N}\right)^2 \partial_z \psi^I(z^\pm) = 0. \quad (6.45)$$

Approximation B assumes that one can solve the surface problem in (6.42) and (6.43) without resorting to truncated series. For instance, with constant or exponential stratifications one can find closed-form, exact expressions for ψ^S [Tulloch and Smith 2009; LaCasce 2012]. Approximation B requires that the two unknown Dirichlet boundary-condition functions $\psi^{S\pm} = \psi^S(z^\pm)$ can be obtained efficiently from specified Neumann boundary-condition functions ϑ^+ and ϑ^- . The Eady problem, discussed below in section 6.5, is a prime example in which one can obtain this Neumann-to-Dirichlet map.

Once ψ^S is in hand, the approximate streamfunction is

$$\psi_N^B = \psi_N^B + \psi^S, \quad (6.46)$$

where the approximate interior streamfunction ψ_N^B is obtained by solving the interior inversion problem (6.44) with the right hand side replaced by the N-mode Galerkin approximation q_N^G defined in (6.17) and (6.18). The two-mode two-surface model of Tulloch & Smith (2009) is the case $N = 1$. The exact solution of the approximate interior inversion problem is

$$\psi_N^B = \sum_{n=0}^N \check{\psi}_n^I(x, y, t) p_n(z), \quad (6.47)$$

where

$$\check{\psi}_n^I \stackrel{\text{def}}{=} \frac{1}{h} \int_{z^-}^{z^+} p_n \psi^I dz, \quad \text{and} \quad \Delta_n \check{\psi}_n^I = \check{q}_n. \quad (6.48)$$

To obtain the approximation B evolution equations we introduce the streamfunction (6.46) into the QGPV equation (6.6) and project onto mode n to obtain

$$\begin{aligned} \partial_t \Delta_n \check{\psi}_n + \sum_{m=0}^N \sum_{s=0}^N \Xi_{nms} J(\check{\psi}_m^I, \Delta_s \check{\psi}_s^I) + \beta \partial_x (\check{\psi}_n^I + \check{\psi}_n^S) \\ + \sum_{s=0}^N \frac{1}{h} \int_{z^-}^{z^+} p_n p_s J(\psi^S, \Delta_s \check{\psi}_s^I) dz = 0, \end{aligned} \quad (6.49)$$

with Ξ_{nms} defined in (6.37). Approximation B assumes that the remaining integral on the second line of (6.49) can be evaluated exactly. This is only possible for particular models of the $N(z)$ (e.g., constant buoyancy-frequency profiles). In practice, however, it may suffice to compute the integral on the second line (6.49) very accurately, e.g. using Gaussian quadrature.

The evolution equations for approximation B are completed with the addition of buoyancy-advection at the surfaces

$$\partial_t \vartheta^\pm + J(\psi^{S\pm}, \vartheta^\pm) + \sum_{n=0}^N p_n^\pm J(\check{\psi}_n^I, \vartheta^\pm) = 0. \quad (6.50)$$

With (6.49) and (6.50) we have $N+3$ evolution equations for the $N+3$ fields $\check{\psi}_0^I, \check{\psi}_1^I, \dots, \check{\psi}_N^I$ and ϑ^\pm .

Approximation B conserves surface buoyancy variance. But the conservation laws for energy is problematic: because ψ^I is not orthogonal to ψ^S the energy (6.10) is not conserved in approximation B (K.S. Smith personal communication). These non-conservative effects are quantitatively small, but are nonetheless irritating. The non-orthogonality of ψ^I and ψ^S was a motivation for development of the surface-aware modes by Smith & Vanneste (2013).

With $\beta = 0$, approximation B conserves potential enstrophy

$$\frac{d}{dt} \sum_{n=0}^N \int \frac{1}{2} (\Delta_n \check{\psi}_n^I)^2 dS = 0. \quad (6.51)$$

But with $\beta \neq 0$ the analog of the exact potential enstrophy (6.12) is not conserved.

6.4.3 Approximation C

Approximation C uses truncated Galerkin approximations ψ_N^G and q_N^G for both ψ and q . The series q_N^G is *not* obtained by differentiation of ψ_N^G and therefore, as indicated in (6.24), the inversion equation is not satisfied exactly by ψ_N^G and q_N^G . But instead, one will have true least-squares approximations to both ψ and q . To our knowledge approximation C, correctly accounting for the surface-buoyancy boundary terms, has not been previously investigated.

Because method C approximates *both* the streamfunction and the PV by Galerkin series, the derivation of the modal equations is very straightforward compared with

the calculations in appendix 6.A of Flierl (1978): one simply substitutes the truncated Galerkin series for the streamfunction (6.15) and PV (6.17) into the QGPV equation (6.6), and then projects onto mode n to obtain

$$\partial_t \check{q}_n + \sum_{m=0}^N \sum_{s=0}^N \Xi_{nms} J(\check{\psi}_m, \check{q}_s) + \beta \partial_x \check{\psi}_n = 0, \quad (6.52)$$

where Ξ_{nms} is defined in (6.37), and we recall the relation between $\check{\psi}_n$ and \check{q}_n from (6.22)

$$\check{q}_n = \Delta_n \check{\psi}_n + \frac{1}{h} (p_n^+ \vartheta^+ - p_n^- \vartheta^-). \quad (6.53)$$

In approximation C there are $N + 3$ degrees of freedom: the $N + 1$ modal amplitudes $\check{\psi}_n$ and the two surface buoyancy fields ϑ^\pm . The approximation C evolution equations are completed by advection of the surface buoyancy

$$\partial_t \vartheta^\pm + \sum_{n=0}^N p_n^\pm J(\check{\psi}_n, \vartheta^\pm) = 0. \quad (6.54)$$

We emphasize that in approximation C the surface buoyancy fields ϑ^\pm are not passive: $\check{\psi}_n$, \check{q}_n , and ϑ^\pm are related through (6.53).

Approximation C conserves surface buoyancy variance as in (6.13). Total energy is also conserved

$$\frac{d}{dt} \sum_{n=0}^N \int \frac{1}{2} |\nabla \check{\psi}_n|^2 + \frac{1}{2} \kappa_n^2 \check{\psi}_n^2 dS = 0. \quad (6.55)$$

With $\beta = 0$, approximation C has a potential enstrophy conservation law

$$\frac{d}{dt} \sum_{n=0}^N \int \frac{1}{2} \check{q}_n^2 dS = 0. \quad (6.56)$$

But with $\beta \neq 0$, as in B, approximation C does not conserve the analog of the exact potential enstrophy (6.12).

6.5 The Eady problem

We use classical linear stability problems with nonzero surface buoyancy to illustrate how solutions to specific problems can be constructed and to assess the relative

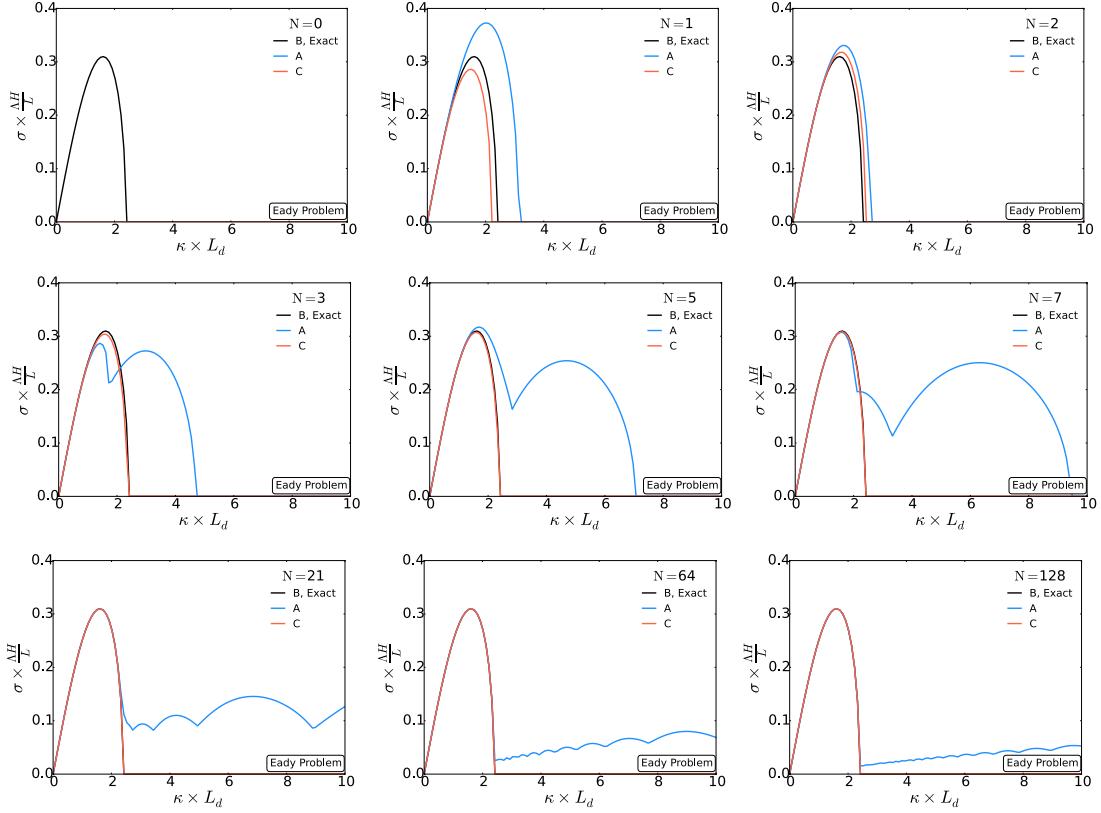


Figure 6.2: Growth rate for the Eady problem as a function of the zonal wavenumber ($l = 0$) using approximations A, B (exact), C with various number of baroclinic modes (N).

merit and efficiency of approximations A, B, and C. The linear analysis does not provide the full picture of convergence of the approximate solutions. Nonetheless, in turbulence simulations forced by baroclinic instability, it is necessary (but not sufficient) to accurately capture the linear stability properties.

We use nondimensional variables so that the surfaces are at $z^+ = 0$ and $z^- = -1$. The Eady exact base-state velocity is given by (6.30) with zero PV $q = 0$ and $\beta = 0$.

6.5.1 Approximation A

While the surface fields ϑ^\pm are dynamically passive in approximation A, the Eady problem can still be considered because the base-state PV defined via (6.35) converges to δ -distributions on the boundaries (Section 6.3).

The base-state velocity in Approximation A is given by the series (6.31) and is a good approximation to the exact base-state velocity (6.30). But, according to approximation A, there is a nonzero interior base-state PV gradient given by the series (6.33). As $N \rightarrow \infty$ the PV gradient in (6.33) converges in a distributional sense to Brethertonian sheets at $z = 0$ and -1 . But for numerical implementation of approximation A we stop short of $N = \infty$. While the PV gradient is much larger at the boundaries, there is always interior structure in the PV (Figure 6.1c). We show that this spurious interior PV gradient has a strong and unpleasant effect on the approximate solution of the Eady stability problem.

To solve the Eady linear stability we linearize the interior equations (6.36) about the base-state velocity in (6.31) and the PV gradient in (6.33). We assume $\check{q}_k = \hat{q}_k \exp[i(kx + ly - \omega^A t)]$, etc, to obtain a $(N + 1) \times (N + 1)$ eigenproblem

$$\sum_{m=0}^N \sum_{s=0}^N \Xi_{nms} [\check{U}_m \hat{q}_s + \partial_y \check{Q}_s \hat{\psi}_m] = c^A \hat{q}_n, \quad (6.57)$$

where \check{Q}_s are the coefficients of the series (6.33) and $c^A \stackrel{\text{def}}{=} \omega^A/k$. The eigenproblem (6.57) can be recast in the matrix form $A\mathbf{q} = c^A \mathbf{q}$, where $\check{\mathbf{q}} = [\hat{q}_0, \hat{q}_1, \dots, \hat{q}_{N-1}, \hat{q}_N]^T$ (appendix 6.B) and solved with standard methods.

Figure 6.2 shows the growth rate of the Eady instability according to approximation A, and compares this with the exact Eady growth rate. Approximation A does not do well, especially at large wavenumbers. The exact Eady growth rate has a high-wavenumber cut-off. At moderate values of N , such as 3, 5 and 7 approximation A produces unstable ‘‘bubbles’’ of instability at wavenumbers greater than the high-wavenumber cut-off. The growth rates in these bubbles are comparable to the true maximum growth rate. As N increases, the unstable bubbles are replaced by a long tail of unstable modes with a growth rate that slowly increases with κ . These spurious high-wavenumber instabilities are due to the rapidly oscillatory interior PV gradient which supports unphysical critical layers: see Figure 6.3.

6.5.2 Approximation B, the exact solution

In approximation B, the zero PV in the Eady problem implies $\check{q}_n = \check{\psi}_n^I = 0$. The $N + 1$ modal equations (with $\beta = 0$) are trivially satisfied; there is no interior

contribution ($\psi_N^B = 0$). Thus approximation B solves the Eady problem exactly.

Assuming $\psi^S = \hat{\psi}^S(z) \exp[i(kx + ly - \omega^B t)]$, we obtain the solution to the surface streamfunction inversion problem (6.42)-(6.43)

$$\hat{\psi}^S(z) = \frac{\cosh[\kappa(z+1)]}{\kappa \sinh \kappa} \vartheta^+ - \frac{\cosh(\kappa z)}{\kappa \sinh \kappa} \vartheta^-, \quad (6.58)$$

where the magnitude of the wavenumber vector is $\kappa = \sqrt{k^2 + l^2}$. We evaluate the surface streamfunction (6.58) at the boundaries to find the relationship between the streamfunction at the surfaces $\hat{\psi}^{S^\pm}$ and the boundary fields ϑ^\pm :

$$\begin{bmatrix} \hat{\psi}^{S^+} \\ \hat{\psi}^{S^-} \end{bmatrix} = \frac{1}{\kappa} \begin{bmatrix} \coth \kappa & -\operatorname{csch} \kappa \\ \operatorname{csch} \kappa & -\coth \kappa \end{bmatrix} \begin{bmatrix} \hat{\vartheta}^+ \\ \hat{\vartheta}^- \end{bmatrix}, \quad (6.59)$$

The nondimensional linearized boundary conditions (6.50) are

$$\hat{\vartheta}^+ - \hat{\psi}^+ = c^B \hat{\vartheta}^+, \quad \text{and} \quad -\hat{\psi}^- = c^B \hat{\vartheta}^-, \quad (6.60)$$

where $c^B = \omega^B/k$. Using the boundary conditions (6.60) in (6.59) we obtain an eigenvalue problem

$$\underbrace{\frac{1}{\kappa} \begin{bmatrix} \kappa - \coth \kappa & \operatorname{csch} \kappa \\ -\operatorname{csch} \kappa & \coth \kappa \end{bmatrix}}_{\text{def } B} \begin{bmatrix} \hat{\vartheta}^+ \\ \hat{\vartheta}^- \end{bmatrix} = c^B \begin{bmatrix} \hat{\vartheta}^+ \\ \hat{\vartheta}^- \end{bmatrix}. \quad (6.61)$$

The eigenvalues of B are given by the celebrated dispersion relation for the Eady problem [Pedlosky 1987; Vallis 2006]

$$c^B = \frac{1}{2} \pm \frac{1}{\kappa} \left[\left(\frac{\kappa}{2} - \tanh \frac{\kappa}{2} \right) \left(\frac{\kappa}{2} - \coth \frac{\kappa}{2} \right) \right]^{1/2}. \quad (6.62)$$

6.5.3 Approximation C

Approximation C expands both the streamfunction and the PV in standard vertical modes. Thus in the Eady problem the PV is exactly zero, as it should be: $q = \check{q}_n = 0$. Thus approximation C does not have the spurious critical layers that bedevil A. Moreover, in approximation C, the N+1 modal equations (with $\beta = 0$) in (6.52) are trivially satisfied, and the inversion relationship (6.53) provides a simple connection between

the streamfunction and the fields ϑ^\pm . The base velocity for the Eady problem in approximation C is the series in (6.31) (the same as A). From the exact shear at the boundaries we obtain the exact base-state boundary variables

$$\Theta^\pm = -y. \quad (6.63)$$

We linearize the boundary equations (6.54) about the base-state (6.33) and (6.63), to obtain

$$\partial_t \vartheta^\pm + U_N^{G\pm} \partial_x \vartheta^\pm - \sum_{k=0}^N \partial_x \check{\psi}_k p_k^\pm = 0. \quad (6.64)$$

Assuming $\vartheta^\pm = \hat{\vartheta}^\pm \exp[i(kx + ly - \omega^C t)]$, and using the inversion relationship (6.53), we obtain a 2×2 eigenproblem

$$C \begin{bmatrix} \hat{\vartheta}^+ \\ \hat{\vartheta}^- \end{bmatrix} = c^C \begin{bmatrix} \hat{\vartheta}^+ \\ \hat{\vartheta}^- \end{bmatrix}, \quad (6.65)$$

where matrix C is defined in appendix 6.C. It is straightforward to show that c^C converges to the exact eigenspeed. i.e., $c^C \rightarrow c^B$ as $N \rightarrow \infty$ (appendix 6.C). Figure 6.2 shows that approximation C successfully captures the structure of the Eady growth rate even with modest values of N.

6.5.4 Remarks on convergence

The crudest truncation (i.e. $N = 0$) is stable for both approximations A and C (Figure 6.2). With one baroclinic mode ($N = 1$) the growth rates ($\omega_i = k \times \text{Im}\{c\}$) are qualitatively consistent with the exact solution, and the results improve with $N = 2$. With a moderate number of baroclinic modes modes ($N > 2$) approximations A and C converge rapidly to the exact growth rate at wavenumbers less than about 2.2 — see figure 6.2. But surprisingly the convergence of the growth rate at the most unstable mode ($\kappa \approx 1.6$) is faster in approximation A ($\sim N^{-4}$) than in approximation C ($\sim N^{-2}$) — see figure 6.4. However, the convergence in approximation C is uniform: there are no spurious high-wavenumber instabilities.

Figure 6.4 also shows that the approximation A convergence of the growth rate to zero at $\kappa = 8$ is slow ($\sim N^{-1}$). While the growth rate does converge to zero at a fixed

wavenumber, such as $\kappa = 8$, we conjecture that there are always faster growing modes at larger wavenumbers.

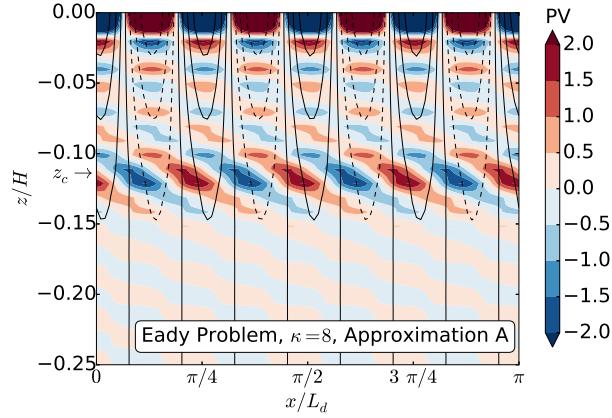


Figure 6.3: Structure of $\kappa = 8$ unstable mode for the Eady problem obtained using approximation A and $N = 64$. Streamfunction is the black curves and PV is the colors. The streamfunction slightly tilts westward as z increases. One can see the unphysical critical layer associated with the fast-oscillating base-state PV. The critical level, z_c , is the depth where the unstable wave speed matches the velocity of the base-state. Only the top quarter of the domain is shown.

6.6 The Green problem

To further explore the relative merit and efficiency of approximations A, B, and C we study the instability properties of a system with nonzero β . For simplicity, we consider a problem with Eady's base-state $\psi = -(1+z)y$ on a β -plane. This is similar to the problem originally considered by Charney [1947] and Green [1960]. The major difference is that Charney considered a vertically semi-infinite domain [Charney 1947; Pedlosky 1987] while we follow Green and consider a finite-depth domain with $-1 < z < 0$.

We obtain the exact system for this “Green problem” by linearizing the QG equations (6.6)-(6.9) about the base-state (6.30) with background PV βy , where $\hat{\beta}$ is the nondimensional planetary PV gradient. Assuming $\psi = \hat{\psi} \exp[i(kx + ly - \omega t)]$, we obtain

$$(U - c)[\hat{\psi}_{zz} - \kappa^2 \hat{\psi}] + \hat{\beta} \hat{\psi} = 0, \quad -1 < z < 0, \quad (6.66)$$

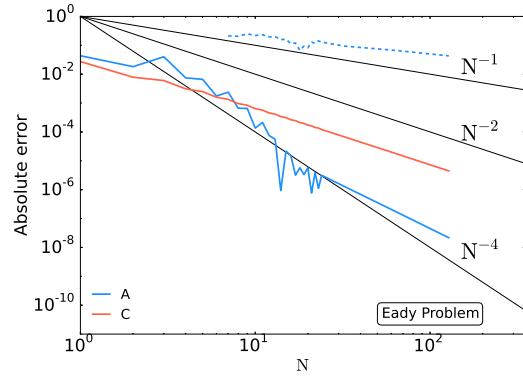


Figure 6.4: Absolute error as a function of number of baroclinic modes (N) for the growth rates of the Eady problem. The solid lines show the error at the exact fastest growing mode ($\kappa \approx 1.6$). The dashed line is the approximation A error at $\kappa = 8$.

and

$$(U - c) \hat{\psi}_z - \hat{\psi} = 0, \quad z = -1, 0. \quad (6.67)$$

As a reference solution, we solve the eigenproblem (6.66)-(6.67) using a centered second-order finite-difference scheme with 1000 vertical levels: see Figure 6.5.

The Green problem supports three classes of unstable modes, indicated in the lower right panel ($N = 128$) of Figure 6.5: (1) the “modified Eady modes”, which are instabilities that arise from the interaction of Eady-like edge waves, only slightly modified by β ; (2) the “Green modes”, which are very long slowly growing modes [Vallis 2006]; (3) the high-wavenumber “Charney modes” are critical layer instabilities that arise from the interaction of the surface edge wave with the interior Rossby wave that is supported by nonzero β .

6.6.1 Implementation of approximation A

The base-state for the Green problem is the same as in the Eady problem. In approximation A, the β -term adds only a diagonal term to the Eady system (6.57) (see appendix 6.C).

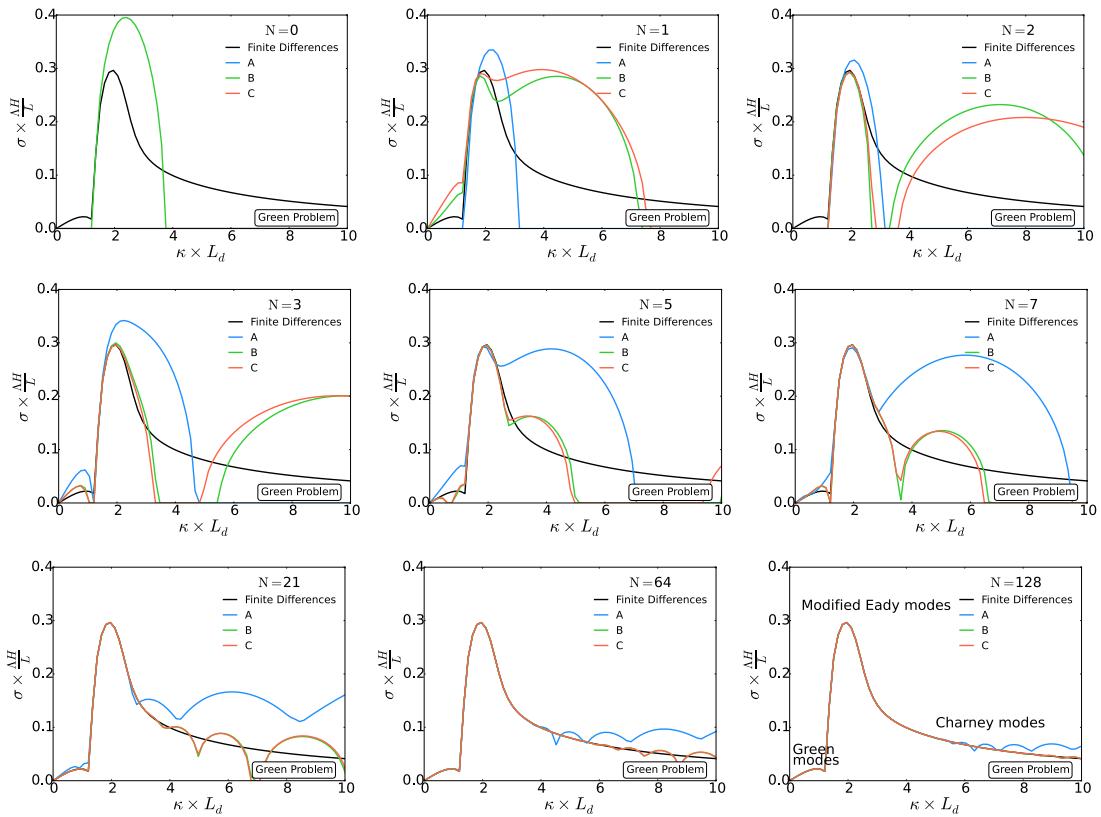


Figure 6.5: Growth rate for the Green problem with $\hat{\beta} = 1$ as a function of the zonal wavenumber ($l = 0$) using approximations A, B, C with various number of baroclinic modes (N). The black line is a finite-differences solution with 1000 vertical levels.

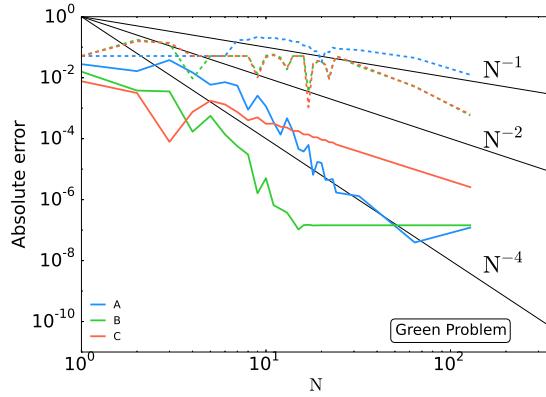


Figure 6.6: Absolute error as a function of number of baroclinic modes (N) for the growth rates of the Green problem. The solid line represent the error at the exact fastest growing mode ($\kappa \approx 1.9$). The dashed line is the error at $\kappa = 8$.

6.6.2 Implementation of approximation B

The base-state is the same as in the Eady problem. The steady streamfunction and buoyancy fields that satisfy (6.49) and (6.50) exactly are

$$\Sigma = -(1+z)y \quad \text{and} \quad \Theta^\pm = -y. \quad (6.68)$$

Assuming $\tilde{q}_n = \hat{q}_n(z) \exp[i(kx + ly - \omega^B t)]$, the $N+1$ interior equations (6.49) linearized about (6.68) are

$$\sum_{s=0}^N \xi_{ns} \hat{q}_s + \hat{\beta} (\hat{\psi}_n + \hat{\psi}_n^S) = c_c^B \hat{q}_n, \quad (6.69)$$

where

$$\xi_{ns} \stackrel{\text{def}}{=} \frac{1}{h} \int_{z^-}^{z^+} p_n p_s (z+1) dz. \quad (6.70)$$

The boundary conditions (6.50), linearized about (6.68), are

$$\hat{\vartheta}^+ - \sum_{s=0}^N p_s^+ \hat{\psi}_s - \hat{\psi}_n^S = c_c^B \hat{\vartheta}^+, \quad (6.71)$$

and

$$-\sum_{s=0}^N p_s^- \hat{\psi}_s - \hat{\psi}_n^S = c_c^B \hat{\vartheta}^-, \quad (6.72)$$

where $\hat{\psi}^S$ is given by (6.58). We use the inversion relationship (6.48) and the Neumann-to-Dirichlet map (6.59) to recast this eigenproblem into standard form $B \tilde{q} = c^B \tilde{q}$, where $\tilde{q} = [\hat{\vartheta}^+, \hat{q}_0, \hat{q}_1, \dots, \hat{q}_{N-1}, \hat{q}_N, \hat{\vartheta}^-]^T$ (see appendix 6.C).

6.6.3 Implementation of approximation C

Again the base-state is the same as in the Eady problem. But now there are $N+3$ equations: the two boundary equations of Eady's problem (6.64) plus $N+1$ interior equations

$$\sum_{m=0}^N \sum_{s=0}^N \Xi_{nms} \check{U}_m \hat{q}_s + \hat{\beta} \hat{\psi}_n = c^C \hat{q}_n, \quad (6.73)$$

We use the inversion relationship (6.53) in (6.73) to recast this eigenproblem in the form $C \tilde{q} = c^C \tilde{q}$, where \tilde{q} is defined as in approximation B (see appendix 6.C).

6.6.4 Remarks on convergence

The most crude truncation ($N = 0$) is stable for approximations A and C. In contrast, the $N = 0$ truncation in approximation B is qualitatively consistent with the modified Eady instabilities: see figure 6.5. With a moderate number of baroclinic modes ($N = 2$ or 3), approximations A, B and C all resolve the modified Eady modes relatively well. At the most unstable modified Eady mode ($\kappa \approx 1.9$), approximation B has typically the smallest error because it solves the surface problem exactly. As in the Eady problem, approximation A converges ($\sim N^{-4}$) faster than approximations B and C ($\sim N^{-2}$) at the most unstable mode, but B and C converge faster at high wavenumbers.

Approximations A, B, and C all converge very slowly to the high-wavenumber Charney modes (Figures 6.5 and 6.6). These modes are interior critical-layer instabilities [Pedlosky 1987] and the critical layer is confined to a small region about the steering level (i.e., the depth at which the phase speed matches the base velocity — see figure ??). With finite base-state shear, the critical layer is always in the interior. Thus, the problem is not that standard vertical modes are inefficient because they do not satisfy inhomogeneous boundary conditions; a low resolution finite-difference solution also presents such “bubbles” in high-wavenumber growth rates (not shown). Resolution of the interior critical layer, not the surface boundary condition, is a problem for

all methods at high wavenumbers. The “surface-aware” modes of Smith and Vanneste [2013] have similar performance to approximations B and C, but also have the same limitation — a large number of vertical modes is required to resolve interior critical layers (K. S. Smith, personal communication).

For example, with $N < 25$, at $\kappa = 8$, approximations are qualitatively inconsistent with the high-resolution finite-difference solution. For larger values of N , the growth rate convergence for approximations B and C scales $\sim N^{-3}$. The growth rate for approximation A converges painfully slowly ($\sim N^{-1}$). As in the Eady problem, at large wavenumbers, the growth rate for approximation A is qualitatively different from that of the finite-difference solution because of spurious instabilities associated with the rapidly oscillatory base-state PV gradient in (6.33).

6.7 Summary and conclusions

The Galerkin approximations A, B, and C are equivalent if there are no buoyancy variations at the surfaces. Thus all three approximations are well-suited for applications with zero surface buoyancy [Flierl 1978; Fu and Flierl 1980; Hua and Haidvogel 1986]. But with nonzero surface buoyancy the three approximations are fundamentally different. In particular, approximation A, originally introduced by Flierl [1978] in a context without surface buoyancy, obtains the approximate PV by differentiating the Galerkin series for the streamfunction, and consequently its approximate PV has violent oscillations in the interior. Approximation B represents the PV as a Galerkin series in standard modes and calculates the streamfunction that satisfies the exact inversion problem associated with the approximate PV [Tulloch and Smith 2009]. The inversion relationship is split into surface and interior problems. Because the surface streamfunction projects onto the interior solution the energy is not diagonalized and consequently approximation B has small errors in energy conservation (K.S. Smith personal communication). The surface-aware modes of Smith & Vanneste (2013) correct this problem. Approximation C uses Galerkin series for both streamfunction and PV but does not satisfy the inversion problem exactly. Nevertheless, the Galerkin series for ψ and q converge absolutely and uniformly, and approximation C provides a good

finite truncation of the QG equations that represents surface buoyancy dynamics and also conserves energy.

With nonzero interior PV gradients the convergence of all approximations is slow for the high-wavenumber Charney-type modes. The critical layer associated with these modes spans a very small fraction of the total depth (Figure ??). To accurately resolve these near-singularities at the steering level there is no better solution than having high vertical resolution in the interior.

For problems with nonuniform surface buoyancy and nonzero interior PV gradient, we recommend approximation C for obtaining solutions to the three-dimensional QG equations using standard vertical modes.

The codes that produced the numerical results of this paper, plotting scripts, and supplementary figures are openly available at github.com/crocha700/qg_vertical_modes.

Appendix 6.A: Convergence of Galerkin series in standard modes

Jackson [1914] gives conditions for the uniform convergence of series expansions in eigenfunctions of the Sturm-Liouville eigenproblem

$$\frac{d^2P_n}{dZ^2} + [\rho_n^2 - \Lambda(Z)] P_n = 0, \quad (6.74)$$

defined on the interval $Z \in [0, \pi]$ with boundary conditions

$$P'_n(0) - \gamma_0 P_n(0) = 0, \text{ and } P'_n(\pi) - \gamma_\pi P_n(\pi) = 0 \quad (6.75)$$

where γ_0 and γ_π are real constants of arbitrary sign and ρ_n^2 is the eigenvalue. The equations defining the standard modes (6.1)-(6.2) can be brought to this form using the following Liouville transformation

$$Z(z) = \frac{1}{\bar{Z}} \int_{z^-}^z S(\xi)^{-1/2} d\xi, \text{ with } \bar{Z} \stackrel{\text{def}}{=} \frac{1}{\pi} \int_{z^-}^{z^+} S(\xi)^{-1/2} d\xi, \quad (6.76)$$

and

$$P_n(Z) = S(z)^{1/4} p_n(z), \text{ where } S(z) \stackrel{\text{def}}{=} \frac{f_0^2}{N^2(z)}. \quad (6.77)$$

The eigenvalues are related by $\rho_n = \bar{Z} \kappa_n$ and

$$\Lambda(Z) = \bar{Z}^2 \left[\frac{1}{4} \frac{d^2 S}{dz^2} - \frac{1}{16S} \left(\frac{dS}{dz} \right)^2 \right]. \quad (6.78)$$

The boundary condition for the standard modes (6.2) implies that the transformed modes satisfy (6.75) with

$$\gamma_0 = \frac{4S(z^-)^{1/2}}{\bar{Z} dS(z^-)/dz}, \quad \text{and} \quad \gamma_\pi = -\frac{4S(z^+)^{1/2}}{\bar{Z} dS(z^+)/dz}. \quad (6.79)$$

If $dS/dz = 0$ at a boundary then the appropriate condition at that boundary is $P_n = 0$.

A special case of Theorem I from Jackson [1914] states that the expansion of a function $f(Z)$ as a series in eigenfunctions P_n converges absolutely and uniformly provided that both df/dZ and $d\Lambda/dZ$ are continuous and bounded, regardless of whether or not f satisfies the same boundary conditions as P_n . (The remainder of the theorem concerns the rate of convergence under stronger conditions on ψ and Λ .) The streamfunction, potential vorticity, and buoyancy profiles are typically assumed to be smooth in studies of QG dynamics, which implies that both f and Λ will satisfy the above conditions. Uniform convergence over $Z \in [0, \pi]$ implies uniform convergence over $z \in [z^-, z^+]$.

Appendix 6.B: Derivation of conservation laws for approximation C

To obtain the conservation of energy in approximation C we multiply the modal equations (6.52) by $-\check{\psi}_n$, integrate over the horizontal surface, and sum on n , to obtain

$$\begin{aligned} & \frac{d}{dt} \sum_{n=0}^N \int [(\nabla \check{\psi}_n)^2 + \kappa_n^2 \check{\psi}_n^2] dS \\ & - \sum_{n=0}^N \frac{1}{h} \int \check{\psi}_n \partial_t (\mathbf{p}_n^+ \vartheta^+ - \mathbf{p}_n^- \vartheta^-) dS \\ & + \sum_{n=0}^N \sum_{m=0}^N \sum_{s=0}^N \Xi_{nms} \int \check{\psi}_n J(\check{\psi}_m, \Delta_s \check{\psi}_s) dS = 0, \end{aligned} \quad (6.80)$$

The triple sum term vanishes by the same symmetry arguments used above in approximation A. The term on the second line of (6.80) is also zero: multiply the boundary

conditions (6.54) by $p_n^\pm \check{\psi}_n$ and integrate over the horizontal surface. Thus we obtain the energy conservation law in (6.55).

The analog of the exact potential enstrophy (6.12),

$$Z_n \stackrel{\text{def}}{=} \sum_{n=0}^N \int \frac{\check{q}_n^2}{2} + \check{q}_n \Delta_n \check{\psi}_n dS, \quad (6.81)$$

is only conserved with $N = 0$.

Appendix 6.C: Details of the stability problems

6.7.1 The interaction tensor

Because the standard vertical modes with constant stratification are simple sinusoids (6.29), the interaction coefficients (6.37) can be computed analytically. First we recall that Ξ_{ijk} is fully symmetric. Permuting the indices so that $i \geq j \geq k$ we obtain

$$\Xi_{ijk} = \begin{cases} 1 & i = j, k = 0; \\ \frac{\sqrt{2}}{2} & i = j + k; \\ 0 & \text{otherwise.} \end{cases} \quad (6.82)$$

The second line in (6.82) corrects a factor of $\frac{1}{2}$ missed by Hua and Haidvogel [1986].

Approximation A

Using the symmetry in Ξ_{nms} , and the inversion relation (6.35), we rewrite row $n + 1$ of the linear Green system

$$\sum_{s=0}^N \sum_{m=0}^N \Xi_{nms} (\check{U}_m + \partial_y \check{Q}_m \alpha_s) \hat{q}_s + \hat{\beta} \alpha_n \hat{q}_n = c^A \hat{q}_n, \quad (6.83)$$

where the inverse of the n 'th mode Helmholtz operator in Fourier space is

$$\alpha_n \stackrel{\text{def}}{=} -(\kappa^2 + (n\pi)^2)^{-1}. \quad (6.84)$$

The Eady problem is the special case $\hat{\beta} = 0$. We use a standard eigenvalue-eigenvector algorithm to obtain the approximate eigenspeed c^A .

Approximation B

The Green eigenvalue problem in (6.69) through (6.72) can be recast in the standard form $B\mathbf{q} = c^B \mathbf{q}$, where $\tilde{\mathbf{q}} = [\hat{\vartheta}^+, \hat{q}_0, \hat{q}_1, \dots, \hat{q}_{N-1}, \hat{q}_N, \hat{\vartheta}^-]^T$. The first and last rows of the system stem from the boundary conditions (6.71)-(6.72)

$$\left(1 - \frac{\coth \kappa}{\kappa}\right)\hat{\vartheta}^+ - \sum_{s=0}^N p_s^+ \alpha_s \hat{q}_s - \frac{\csch \kappa}{\kappa} \hat{\vartheta}^- = c^B \hat{\vartheta}^+, \quad (6.85)$$

and

$$\frac{\csch \kappa}{\kappa} \hat{\vartheta}^+ - \sum_{s=0}^N p_s^- \alpha_s \hat{q}_s + \frac{\coth \kappa}{\kappa} \hat{\vartheta}^- = c^B \hat{\vartheta}^-. \quad (6.86)$$

The $(n+1)$ 'th row originates from the n 'th interior equation (6.69)

$$-\hat{\beta} p_n^+ \alpha_n \vartheta^+ + \sum_{s=0}^N \gamma_{ns} \hat{q}_s + (\beta \alpha_n + 1) + \hat{\beta} p_n^- \alpha_n \vartheta^- = c^B \hat{q}_n, \quad (6.87)$$

where the symmetric matrix γ_{ms} is

$$\gamma_{ij} \stackrel{\text{def}}{=} \int_{-1}^0 p_i p_j z dz = \begin{cases} -\frac{1}{2} : & i = j; \\ \frac{2\sqrt{2}}{(j\pi)^2} : & i = 0, j \text{ is odd}; \\ \frac{4(i^2+j^2)}{[(i^2-j^2)\pi]^2} : & i + j \text{ is odd}. \end{cases} \quad (6.88)$$

Approximation C

The Eady problem

The 2×2 eigenproblem is

$$\underbrace{\begin{bmatrix} U_N^{G+} + \Sigma_N & -\Omega_N \\ \Omega_N & U_N^{G-} - \Sigma_N \end{bmatrix}}_{\stackrel{\text{def}}{=} C} \begin{bmatrix} \hat{\vartheta}^+ \\ \hat{\vartheta}^- \end{bmatrix} = c^C \begin{bmatrix} \hat{\vartheta}^+ \\ \hat{\vartheta}^- \end{bmatrix}, \quad (6.89)$$

where

$$\Sigma_N \stackrel{\text{def}}{=} \alpha_0 + 2 \sum_{n=1}^N \alpha_n, \quad \text{and} \quad \Omega_N \stackrel{\text{def}}{=} \alpha_0 + 2 \sum_{n=1}^N (-1)^n \alpha_n. \quad (6.90)$$

The sums (6.90) become exact in the limit $N \rightarrow \infty$

$$\Sigma_\infty = -\frac{\coth \kappa}{\kappa}, \quad \text{and} \quad \Omega_\infty = -\frac{\csch \kappa}{\kappa}. \quad (6.91)$$

The base velocity also converges to the exact result. Using standard results for the summation of inverse squares, we obtain

$$U_{\infty}^G + = 1, \quad \text{and} \quad U_{\infty}^G - = 0. \quad (6.92)$$

Thus

$$C \rightarrow B \quad \text{as} \quad N \rightarrow \infty, \quad (6.93)$$

and the eigenvalues of the Eady problem using approximation C become exact i.e., $c^C \rightarrow c^B$ as $N \rightarrow \infty$.

The Green problem

The $(N+3) \times (N+3)$ eigenproblem is

$$C \tilde{q} = c^C \tilde{q}, \quad (6.94)$$

where \tilde{q} is defined as above in approximation B. The first and last rows of (6.94) stem from the boundary conditions (6.64)

$$(U_N^{G+} + \Sigma_N) \hat{\vartheta}^+ - \sum_{n=0}^N \alpha_n p_n^+ \hat{q}_n - \Omega_N \hat{\vartheta}^- = c^C \hat{\vartheta}^+, \quad (6.95)$$

and

$$\Omega_N \hat{\vartheta}^+ - \sum_{n=0}^N \alpha_n p_n^- \hat{q}_n + (U_N^{G-} - \Sigma_N) \hat{\vartheta}^- = c^C \hat{\vartheta}^-. \quad (6.96)$$

Row $n+1$ originates from the n 'th modal equation (6.73):

$$\begin{aligned} \hat{\beta} \alpha_n p_n^+ \hat{\vartheta}^+ + \sum_{s=0}^N \sum_{m=0}^N \Xi_{mns} \check{U}_m \hat{q}_s + \hat{\beta} \alpha_n \hat{q}_n \\ - \hat{\beta} \alpha_n p_n^- \hat{\vartheta}^- = c^C \hat{q}_n. \end{aligned} \quad (6.97)$$

Acknowledgments

CBR is grateful for a helpful conversation with G. R. Flierl. We thank Joe LaCasce and Shafer Smith for reviewing this paper. We had useful discussions with Shafer Smith regarding approximation B and the “surface-aware” modes, and with Geoff Vallis — who

pointed out that Green (1960) first considered the Eady+ β problem. This research was supported by the National Science Foundation under award OCE 1357047.

Chapter 6, in full, reprints material as it appears in Journal of Physical Oceanography, 2016, 46, doi:10.1175/JPO-D-15-0073.1. Rocha, C. B.; Young, W. R.; Grooms, I. The dissertation author was the primary investigator and author of this paper.